

Immunohistologisten värjäysten digitaalinen analysointi – katsaus vaihtoehtoihin ja toteutus käyttäen QuPath-ohjelmistoa

Juha Hämäläinen

Diplomi-insinööri, hammaslääketieteen kandidaatti

Suu- ja leukasairauksien osasto

Helsinki 11.9.2018

Tutkielma

juha.hamalainen@helsinki.fi

Ohjaaja: Tuula Salo

HELSINGIN YLIOPISTO

Lääketieteellinen tiedekunta

HELSINGIN YLIOPISTO – HELSINGFORS UNIVERSITET

| | | | |
|--|--|---|--|
| Tiedekunta/Osasto – Fakultet/Sektion – Faculty Suu- ja leukasairauksien osasto | | Laitos – Institution – Department Suupatologia | |
| Tekijä – Författare – Author Juha Hämäläinen | | | |
| Työn nimi – Arbetets titel – Title Immunohistologisten värjäysten digitaalinen analysointi – katsaus vaihtoehtoihin ja toteutus käyttäen QuPath-ohjelmistoa | | | |
| Oppiaine – Läroämne – Subject Suupatologia | | | |
| Työn laji – Arbetets art – Level Tutkielma | Aika – Datum – Month and year 11.9.2018 | Sivumäärä - Sidoantal - Number of pages 14+34 | |
| <p>Tiivistelmä – Referat – Abstract</p> <p>Digitaalisella analyysillä tarkoitan tässä yhteydessä kudoksenäytteiden analysoimista tietokoneavusteisesti: immunohistologisesti värjäyty näytelasi skannataan ja tiedosto avataan tietokoneelle ladatussa ohjelmistossa, jossa se analysoidaan. Ohjelma tunnistaa soluja sekä värisävyeroja. Tutkijat haluavat mennä kohti digitaalista analyysia, koska sen uskotaan olevan objektiivisempaa ja tehokkaampaa kuin manuaalisen visuaalisen analyysin, varsinkin suurien aineistojen kohdalla.</p> <p>Tutkimuksessa selvitin immunohistologisten näytteiden digitaalisen analysoinnin nykytilaa ja mahdollisuuksia tulevaisuudessa. Tavoitteenani oli saada kuva digitaalisen analysoinnin käytöstä, ja kirjallisuuskatsauksen sekä haastattelujen perusteella valita analysointiohjelmisto suupatologian tutkimusryhmän käyttöön. Tutkimuksessa haastattelin tutkijoita Oulun ja Helsingin yliopistoista, osallistuin aiheita käsitelleeseen kahteen seminaariin, laadin kirjallisuuskatsauksen sekä perehdyin ohjelmistojen käyttöön.</p> <p>Ohjelmistojen käyttö on kansainvälisesti laajaa, ja tutkijakunnassa suositaan kaupallisten ohjelmien sijaan nk. <i>open-source</i> ilmaisohjelmistoja, kuten ImageJtä, CellProfileriä tai Icyä. Haastattelujen ja kirjallisuuskatsauksen perusteella valikoin testikäyttöön QuPath-ohjelmiston. Testasin ohjelmiston toimivuutta vertaamalla digitaalista analyysia tutkijoiden silmämääräisesti tekemään immunohistologiseen analyysiin kymmenestä kielisyöpänäytteestä.</p> | | | |

HELSINGIN YLIOPISTO – HELSINGFORS UNIVERSITET

Tiivistelmä – Referat – Abstract

Tein ohjelmalle seikkaperäiset analysointiohjeet suomeksi ja englanniksi. Kolme tutkijaa käytti ohjeideni mukaan ohjelmaa ja antoi siitä kommenttinsa. Ohjelman käyttö oli heidän mielestään helppoa, mutta sen avulla kokonaisten kudosnäytteiden analyysi oli hankala: tutkijan tuli valita analysoitava alue, mikä saattoi vaikuttaa tuloksiin. Tutkijoiden mukaan ohjelma sopii hyvin *mikroarray*- tyyppisiin, pelkästään kasvainkudoksen soluja sisältävien kudosnäytteiden immunovärjäyksien analyysihin. Sen sijaan sen käyttö on hankalampaa kokonaisissa kudosleikkeissä, joissa antigeenin sijainnissa on gradienttiero eri kasvainvyöhykkeiden alueilla. Se ei myöskään sovellu tapauksiin, joissa tulisi analysoida antigeenin tarkka sijainti suhteessa soluun.

Avainsanat – Nyckelord – Keywords

Pathology; Immunohistochemistry, Computer-assisted Image Analysis; QuPath

Säilytyspaikka – Förvaringställe – Where deposited

Tiedekunnan kanslia toimittaa Terkkoon. Opiskelija tallettaa sähköisen version Heldaan.

Muita tietoja – Övriga uppgifter – Additional information

| | | |
|-----|---|----|
| 1 | Johdanto | 1 |
| 2 | Kuivanalysointiohjelmat | 1 |
| 2.1 | ImageJ..... | 2 |
| 2.2 | Fiji..... | 3 |
| 2.3 | Icy..... | 3 |
| 2.4 | QuPath | 4 |
| 2.5 | CellProfiler | 4 |
| 2.6 | Imaris..... | 5 |
| 3 | Digitaalisen analysoinnin käyttö | 6 |
| 4 | Tutkimuksen tavoitteet..... | 7 |
| 5 | Tutkimusaineisto ja – menetelmät | 7 |
| 6 | Analysoinnin toteutus QuPath-ohjelmistolla..... | 10 |
| 7 | Tulokset | 11 |
| 7.1 | Vertailu manuaaliseen analyysiin | 11 |
| 7.2 | Käyttäjien kokemukset..... | 11 |
| 8 | Pohdintaa..... | 12 |
| 9 | Kiitokset | 14 |
| | Lähdeluettelo..... | 15 |
| | Liite 1: QuPath – Analysointiohje..... | 18 |
| | Liite 2: QuPath – Analysis Instructions (engl.) | 37 |

1 Johdanto

Tutkimus käsitteli immunohistologisten näytteiden värjäytymisen analysoinnin siirtämistä digitaalisesti tehtäväksi. Nykyisellään Helsingin yliopiston suu-patologian tutkimuksessa värjäytyminen analysoidaan aistinvaraisesti silmillä mikroskoopista katsoen kahden tai tarvittaessa kolmen tutkijan tekemänä.

Viime vuosina näytteiden digitaalinen analysointi on tullut yleisemmäksi ja siitä on julkaistu tieteellisiä artikkeleita. Analysoinnin tueksi on kehitetty sekä kaupallisia tuotteita että ilmaisohjelmistoja. Digitaalisen analysoinnin hyödyiksi voidaan alustavasti lukea objektiivisuuden, paremman toistettavuuden ja suuremman tehokkuuden. Lisäksi digitaalisesti tehty näytteiden analysointi on etu, kun artikkeleita pyritään saamaan tieteellisiin julkaisuihin julkaistaviksi.

Tutkimusta aloittaessa oli tiedossa, että Helsingin ja Oulun yliopistoissa on käytetty eri tutkimusryhmissä eri analysointiohjelmistoja kokeiluluonteisesti tai yksittäisten tutkijoiden toimesta, mutta laajempaa digianalytiikan käyttö ei vielä ole. Tutkimuksessa tehtiin aiheesta kirjallisuuskatsaus ja haastatteluita, osallistuttiin aihetta käsitteleviin seminaareihin ja näiden synteesinä valittiin yksi ohjelmisto testikäyttöön. Tämän ohjelmiston analysointikäyttöön laadittiin seikkaperäiset ohjeet.

Tulevissa luvuissa on käyty läpi tämä prosessi – aluksi esitellään otos kuvan-analysointiohjelmistoja (luku 2) kirjallisuuskatsauksen pohjalta ja tätä seuraa katsaus digianalyysin nykytilaan (luku 3). Tutkimuksen eri tavoitteet on kuvattu luvussa 4 ja käytettävissä olleet lähteet ja menetelmät luvussa 5. Tutkielma jatkuu analysoinnin tarkempaan esittelyyn valitulla ohjelmistolla (luku 6) ja päättävinä lukuina esitetään tulokset (luku 7) sekä suosituksia tulevaan antava pohdintaa (luku 8).

2 Kuvananalysointiohjelmat

Kuvananalysointiin on kehitetty luonnollisesti kaupallisia sovelluksia, mutta ilahduttavan paljon myös vapaan lähdekoodin pohjalta toimivia ilmais-

ohjelmistoja. Ilmaisohjelmistot tuntuvat olevan vieläpä varsin suosittuja tutkimusryhmissä, joten katsauksessa keskitytään enemmän niihin. Ilmaisohjelmistoista esitellään ImageJ, Fiji, Icy, CellProfiler sekä QuPath ja kaupalliselta puolelta Imaris, joka on käytössä myös Helsingin Yliopiston tietyllä laboratoriollla.

2.1 ImageJ

Yksi paljon käytetty ohjelmisto on ImageJ. ImageJ on ilmaisohjelmisto ja se on kehitetty Yhdysvaltain National Institute of Healthin työntekijän toimesta. Se on Java-pohjaisena käytettävissä kaikilla yleisillä käyttöjärjestelmillä. Yhtenä vahvuutena ohjelmistolla on runsas käyttäjäkunta – esimerkiksi yli 1700 käyttäjää ja kehittäjää postituslistalla – ja runsas plugin-tarjonta. Plugineilla ohjelmistoa saadaan hyödynnettyä ja suunnattua moniin erikoistarpeisiin, ja esimerkiksi solujen laskentaan on julkaistu plugineja. (1,2)

ImageJ ohjelmistoa on hyödynnetty esimerkiksi Oulun Yliopistossa tehdyssä tutkimuksessa paksusuolen syövän näytteiden analysoinnista digitaalisesti, missä se osoitti toimivansa hyvin ja tarjoavan varteenotettavan vaihtoehdon manuaaliselle ja silmämääräiselle näytteiden analysoinnille. (3) Ohjelman käytöstä lääketieteellisiin tarpeisiin on julkaistu ja koko ajan julkaistaan useita artikkeleita. (4-7) Sen onkin tunnustettu olevan pioneeri digitaalisten analysointi-ohjelmistojen saralla (8). Alkuperäisen ImageJ:n suosion myötä myös sen rajoitteita on tunnistettu ja vuonna 2017 siitä julkaistaan kokonaan uudelleenkirjoitettu versio ImageJ2 (9).

Vuonna 2015 tutkijaryhmä teki analysointiohjelmistojen vertailun käyttäessä fluorosenssimikroskooppikuvia. Tässä vertailussa arvioitiin ohjelmistoja sekä käytettävyyden että toiminnallisuuksien laajuuden puolesta. (10) Arvioinnissa olivat mukana tässä tutkielmassa esitellyistä ImageJ, Fiji, Icy sekä CellProfiler. ImageJ jäi käytettävyydessä viimeiseksi, mutta oli toiminnallisuuksien laajuudessa miltei parhaat pisteet saaneen CellProfilerin rinnalla. (10)

2.2 Fiji

ImageJ:n avoin lähdekoodi ja runsas plugin-tarjonta mahdollistavat hyvän lähtökohdan jatkokehitykselle spesifisiin tarpeisiin. Fiji-ohjelmisto, tai pikemminkin plugin-paketti on juuri tältä pohjalta kehitetty. Se oli saavuttanut jo vuonna 2012 n. 20 000 käyttäjää. (8,11)

Fiji kokoaa yhteen noin 200 kappaletta plugineita, jotka kehittäjät ovat nähneet hyödylliseksi tieteellisessä kuva-analyysissä. Kuten alla esitellyssä Icy:ssä, myös Fijin kehittäjät toivovat siitä yhteistä alustaa tietojenkäsittelytieteilijöiden ja biologisten välille. Toisaalta Fijin pyrkimyksenä on olla muillekin kuin ohjelmoijille helposti lähestyttävä, eli tutkijoiden pitäisi pystyä toteuttamaan sillä helposti vaativiakin kuvankäsittelyalgoritmeja ja -scriptejä. (11)

Fijin etuna ImageJ:hin ovat päivitetty ja helpotettu arkkitehtuuri uusien algoritmien luomiseen useilla eri ohjelmistokielillä. Lisäksi analysointipluginien päivityksestä ja jakelusta on tehty tehokasta. Fiji integroituu moniin muihin ohjelmistoihin, kuten Matlabiin tai CellProfileriin. Fijin kehityksessä on otettu huomioon myös ImageJn uusien kehitys ImageJ2:ksi ja Fijin pääkirjastot ja -pluginit tukevat myös ImageJ2:sta. (11)

Vuoden 2015 vertailussa Fiji oli rinnastettu ImageJ:hin, joten erillisiä tuloksia sen käytettävyydestä tai toiminnollisuuksista muihin verrattuna ei tästä tutkimuksesta saatu. (10)

2.3 Icy

Icy on yksi laajemmin käytössä olevista avoimen lähdekoodin kuvien analysointiohjelmisto. Se on kehitetty Pasteur Instituutissa. Tarkemmin ottaen se tarjoaa joustavan alustan, jolla toisaalta matemaatikot pääsevät luomaan algoritmeja ja jolla biologit tai muut tutkijat analysoivat kuvia. (12)

Icy:n ominaisuuksia ovat muun muassa javascript tai python -pohjaiset scriptaukset, ImageJ-integraatio sekä arvosteltavat ja kommentoitavat pluginit. (13)

Icy:n kehitys lähti ImageJ:n huomatuista puutteista, ja se pyrkii olemaan parempi muun muassa joustavuuden, datan visualisoinnin ja kehitettävyyden kannalta. Biologeille, ja muille tutkijoille, se tarjoaa helposti selattavan plugin-kirjaston ja tarkoitukseen sopivan analysointiaplikaation. Kehittäjille korostettavia asioita ovat muun muassa kehitysystävällinen ympäristö sekä tehokkaat jakelu- ja palautekanavat. (12)

Vuonna 2015 tehdyssä vertailussa Icy oli käytettävyydessä ImageJta parempi, mutta CellProfileria heikompi. Toiminnollisuuksien laajuudessa se jäi kolmikon viimeiseksi. (10)

2.4 QuPath

QuPath on löyhästi nimensä mukaisesti kehitetty nimenomaan digitaalisen patologian tarpeisiin. Sekin on vapaan lähdekoodin ohjelmisto ja se on kehitetty Queensin yliopistossa Belfastissa ja sen Centre for Cancer Research & Cell Biology yksikössä.

Kyseessä on tuore analysointiohjelmisto, joka pyrkii auttamaan digitaalisessa patologiassa etenkin erittäin suurten 2D-kuvien ja whole slide image (WSI) analyysissä. Edellä mainituille analysointitarpeille ei ole ollut muissa aiemmin kehitetyissä ilmaisohjelmoistoissa kunnollista tukea, ja tämä osaltaan johti QuPathin kehittämiseen. Sen toiminnollisuuksiin kuuluvat tuumorien ja biomarkkerien tunnistustyökalut sekä batchien ja scriptien ohjelmointimahdollisuudet, joilla tutkijat voivat tehostaa näytteiden analysointia. Lisäksi se on yhteensopiva monien muiden analysointiohjelmistojen kanssa, mikä mahdollistaa niiden vahvuuksien hyödyntämisen QuPathin rinnalla. (14)

2.5 CellProfiler

Samoin kuin ImageJ ja muut yllä esitellyt ohjelmistot, CellProfiler on myös open-source, eli vapaan lähdekoodin ilmaisohjelmisto. CellProfiler on alun pitäen suunniteltu jo hivenen nimestä päätellenkin biologisten analysointitarpeisiin. Sen kehityksestä vastaa Harvardin ja MIT:n Broad Instituten alainen Carpenter Labissa työskentelevä tiimi. Tiimi korostaa CellProfilerin helppokäyttöisyyttä

projektin internet-sivuilla mainitsemalla, että ohjelmiston käyttö ei vaadi aiempaa kokemusta ohjelmoinnista eikä konenäöstä. (15)

CellProfiler pystyy käsittelemään kaikkia tyypillisimpiä mikroskooppikuvia, ja sen ominaisuuksiin kuuluu muun muassa sekä analysointia edeltävä muokkaus, kuten tiettyjen haluttujen kuvan piirteiden vahventaminen, että objektien tunnistus käyttäen koneoppimista (engl. machine learning) tai kuvien binäärisiksi muuttamista ja sitä kautta tunnistuksen helpottamista. (16)

Ensimmäinen versio CellProfilerista esiteltiin jo vuonna 2005 ja nyt tuorein esitelty sukupolvi on 3.0, jossa on tuotu mukaan mm. 3D-kuvien analysointia ja pilviominaisuuksia. Ohjelman hyödyllinen piirre läpi eri sukupolvien on ollut ja on haluttujen muokkauspolkujen (engl. macro tai pipeline) luominen, joissa siis halutut muokkaukset ja analyysit saadaan toteutettua aina uusille näytteille automatisoituna työjonona, mikä nopeuttaa analysointia ja mahdollistaa suurten näytemassojen käsittelyn. (17)

CellProfileria on käytetty ja käsitelty myös useasti viime aikaisissa julkaisuissa koskien digitaalisen analysoinnin hyödyntämistä lääketieteessä. (18,19) Vertailussa ImageJn ja Cellprofilerin välillä korostuu se, että ImageJ pyrkii olevamaan joustava analysointiohjelmisto, kun taas CellProfilerin vahvuus ovat peräkkäisten työvaiheiden käsittely suurillekin kuvajoukoille tehokkaasti (9).

Vuoden 2015 vertailussa CellProfiler sai parhaat pisteet sekä käytettävyydestä että toiminnollisuuksien laajuudesta. Haasteena on kuitenkin mainittu jo ”runsauden pula”, eli käyttäjät eivät välttämättä osaa valita optimaalisinta analysointitapaa suuresta tarjonnasta, mikä vaatisi enemmän kuvankäsittelyn osaamista. (10)

2.6 Imaris

Imaris on yksi tunnetuista, ja Helsingin yliopistollakin käytössä oleva, kaupallinen mikroskooppikuvien käsittelyyn ja analysointiin kehitetty ohjelmisto. Imarixen kehityksestä ja myynnistä vastaa yhtiö nimeltä Bitplane, mutta alun perin tämänkin ohjelmiston juuret ovat tiedemaailmassa ja tarkemmin Zurichin teknillisessä yliopistossa (Eidgenössische Technische Hochschule Zürich). (20)

Imaris on esitelty jo vuonna 1993 ja se on alusta pitäen suunniteltu biotieteiden tarpeita ajatellen. Sen ominaisuuksiin kuuluvat mm. 3D- ja 4D-datasettien käsittely, vahvat visualisointi- ja merkitsemisominaisuudet sekä modulaarinen rakenne, eli ohjelmiston voi ostaa vain tutkijalle tarpeellisilla analysointi-moduuleilla. Toisaalta siinä on myös integraatiomahdollisuus ImageJ:n/Fijiin plugin-kirjastoihin. (20)

Ohjelmiston hintoja ei ole julkaistu nettisivuilla, vaan ne pitäisi tiedustella valmistajalta. Nettihakujen perusteella kuitenkin viitteelliseksi vuosihinnoiksi voidaan arvioida useita tuhansia euroja per moduuli.

Jos ajatellaan kaupallisten ohjelmistojen rajoitteita, niin hinnan lisäksi niiden analysointialgoritmit on usein piilotettu ja käyttäjät eivät täten pääse niitä myöskään räätälöimään, mikä on nähty puutteeksi tutkijakunnassa. (11)

3 Digitaalisen analysoinnin käyttö

Digitaalisen analysoinnin käytöstä on raportteja jo 1980-luvulta alkaen, mutta näihin päiviin asti tällaisten laitteiden tai ohjelmistojen saatavuutta on haitannut niitten korkea hinta. Viime aikoina saataville tulleet ilmaisohjelmistot ovatkin tarjonneet houkuttelevan ja käyttökelpoisen vaihtoehdon immunohistokemiallisten näytteiden analysointiin. (3)

Tieteellisiä artikkeleita, joissa ohjelmistoja käytetään analysoinnin apuna, on julkaistu todella runsaasti. Esimerkkinä hakusanajono "digital image analysis" tuottaa pelkästään pubmedissä yli 10 000 osumaa. Asiaa voi lähestyä myös viittausmäärien kautta, ja esimerkiksi CellProfileriin on viitattu yli 6000 kertaa, ja tämän hetkinen viittaustahti on yli 1000 per vuosi (17). Ohjelmistojen kehittäjät seuraavat myös viittauksia tahoillaan ja esimerkiksi QuPathin kotisivuilla listataan 37 QuPathia hyödyntänyttä tieteellistä artikkelia vuosilta 2017-2018.

Esimerkkejä sovellutuksista on siis todella runsaasti. Tutkimusryhmät ovat esimerkiksi kehittäneet makroja ImageJ:n päälle ja ratkaisseet näille omia tutkimusongelmiaan, kuten immunofluorosenssikuvista eri parametrien tunnistamisen hermokudoksessa (4), solujen ja soluyhdyskuntien automaattisen laskennan paremmuuden arvioinnin manuaaliseen verrattuna (21) tai

valitun ohjelmiston testauksen tutkimusaineistona käytettiin kymmentä immunohistologisesti värjättyä näytettä.

Tutkimusmenetelminä vastaavasti käytettiin ensinnäkin kirjallisuuskatsausta, eli tutustuttiin viime aikoina aiheesta julkaistuihin artikkeleihin ja pyrittiin muodostamaan kuva sopivimmista ohjelmistoista, niiden käyttötavoista ja parametreista näytteiden analysointiin. Toisekseen tutkimusmenetelmänä käytettiin haastattelututkimusta, ja haastateltaviksi ohjaajan suosituksesta pyydettiin tutkijoita sekä Helsingin että Oulun yliopistoista, joilla tiedettiin olevan kokemusta digitaalisen analysoinnin parista. Suunniteltuja haastateltavia oli 5 kappaletta ja he kaikki suostuivat haastatteluun.

Tohtori Fang Zhao toimii Helsingin yliopiston Haartman-instituutin mikroskopointiyksikössä ja hänen kokemuksensa mukaan ImageJ-ohjelmiston asema on vahva tutkijakentässä osin siksi, että se on pisimpään saatavilla. Kaupalliset ohjelmat ovat osoittautuneet kalliiksi, ja toisaalta ilmaisohjelmistot tarjoavat samat toiminnallisuudet ehkä hieman vaatimattomilla käyttöliittymillä. (2)

FIMMin tutkija, LL Oscar Brück on käyttänyt kolmea eri ohjelmistoa – ImageJ:tä, CellProfileriä ja uusimpana QuPathia. Hän on eri tutkimustarkoitusten mukaan käyttänyt esimerkiksi CellProfileriä solunlaskentaan, kun taas ImageJ on muuhun analysointiin, kuten intensiteetin arviointiin, etevämpi. QuPathiin Brück on myös tutustunut, ja saanut joitain kehittäjälle ehdottamiaan parannuksiakin siihen lisättyä. QuPathin etuna on muun muassa suurten kuvajoukkojen patch-ajot, mutta toisaalta se on nopeasti katsottuna aika kompleksi. Mitä tahansa ohjelmistoa käyttääkään, niin laadukas lähtökohta, eli digitalisoitu näytelasi, on oleellinen. FIMMillä on käytössään oma digitaalinen skannauspalvelu tähän tarkoitukseen. Tieteellisissä artikkeleissa Brückin kokemuksen mukaan on yhä vallalla ”kultainen standardi” kahden patologin arvion käyttämisestä, mutta tämä tulee muuttumaan, joskin hitaasti. (27)

Oulun Yliopistolla on käytetty ilmaisohjelmista sekä ImageJ:tä, ImmunoRatiota että QuPathia ja tämän lisäksi on ollut käytössä kaupallinen ImagePro Analyzer, joka on koettu kalliiksi (28,29). Dosentti Juha Väyrynen on kehittänyt ImageJ-pohjaisen konfiguroitavan makron, jota on käytetty varsinkin tulehdussolujen

laskennassa erityisesti Oulussa, mutta myös Turussa ja Jyväskylässä (28). Digitaalisesta analysoinnista yleisesti nostettiin esiin näytesarjojen tasalaatuisen digitalisoinnin tärkeys sekä toisaalta analysoinnin rajoitteet – tumat on helppo tunnistaa monilla ohjelmilla, mutta degranulaatiot ynnä muut vaativat soveltamista ja ohjelmien syvällisempää käyttöä (28). Patologian erikoislääkäri Juha Näpänkangas on perehtynyt kliiniseen digipatologiaan ja sen osana tai ohella myös digitaalisen kuva-analyysiin. Sekä Näpänkangas että Väyrynen suosittelivat vahvasti QuPathiin perehtymistä ja testianalyysien tekemistä sillä (28,29). Vaikka QuPath oli alkuun varsin raakile, niin nyt sen käyttöliittymä on hyvä ja yhtenä selkeänä etuna ImageJ:hin nähden on skannattujen leikkeiden käsittely kokonaisena, kun taas ImageJ:ssä näiden käsittely on parsimista (29).

Oulussa ja Linköpingissä kliinisessä työssä digitaalisen analysoinnin käyttö on vielä vähäistä, ja digipatologian pääkehitys tällä rintamalla onkin ollut skannauksen ja siihen liittyvän prosessin kehityksessä. Tosin järjestelmätoimittajat ovat kehittämässä analysointiominaisuuksia myös kliinisten sovellusten päälle. (29,30) Linköpingissä kliininen digipatologia on siis sielläkin näytteen tarkastelemista ruudulta ja esimerkiksi merkintöjen tekemistä kuvaan, mutta varsinaista analyysiä ei tietokoneella tai tietokoneen toimesta tehdä. Hekin kuitenkin ovat kehittämässä kuva-analyysiä valitsemansa järjestelmätoimittaja Sectran kanssa yhteistyössä. (30)

Lopuksi tutkimuksessa analysoitiin kymmenen esimerkinäytettä uudella esitetyllä tavalla ja saatuja tuloksia verrattiin samojen näytteiden perinteisen, eli aistinvaraisen, analyysin tuloksiin. Näytteistä analysoitiin värjäytymän intensiteetti. Intensiteetti päädyttiin analysoimaan näytteiden kasvainpesäkkeiden viereisiltä tulehdussolualueilta siten, että kustakin näytteestä valittiin 3 analyysialuetta eri kohdista, ja kun analyysialueen tuli sisältää noin 100 solua. Intensiteetti määritettiin koko solun, eli tuman ja sytoplasman, keskiarvona. Intensiteettiarvot määritettiin kolmeen luokkaan, eli kukin solu sai luokituksen negatiivinen, luokka 1, luokka 2 tai luokka 3 (vahvin värjäytymä). Lopulta kokonaisintensiteettiä kuvaava H-score saatiin laskemalla oheisella kaavalla:

$$H - Score = [1 \times (\% \text{ soluista, luokka 1}) + 2 \times (\% \text{ soluista, luokka 2}) + 3 \times (\% \text{ soluista, luokka 3})]$$

Kaavasta käy ilmi, että H-Score voi saada arvoja väliltä 0-300 ja se toimii intensiteetin ja värjäytyneiden solujen määrän yhteistunnuslukuna.

Roolitus tutkimus jakaantui siten, että opiskelija vastasi kirjallisuuskatsauksesta ja haastatteluista omatoimisesti ohjaajan tukemana. Esimerkinäytteiden analysoinnissa tutkimuslaboratoriosta nimetty tutkija auttoi digitaalisten näytteiden esimerkkialueiden valinnassa ja kaksi muuta tutkijaa olivat tehneet aistinvaraisen analyysin aiemmin.

6 Analysoinnin toteutus QuPath-ohjelmistolla

Tehtyjen haastattelujen, kirjallisuuskatsauksen ja ohjelmistoihin tutustumisen perusteella testikäyttöön valittiin QuPath -ohjelmisto. Näytteiksi valittiin kielisyöpänäytteitä, jotka olivat immunohistokemiallisesti värjätty nikotiini-reseptori-alfa-7 osalta. Näytelasit skannattiin Oulussa 5.9.2017 ja kuvatiedostot saatiin analysointiin siirrettävällä kovalevyllä. On huomattavaa, että skannattujen näytteiden kuvatiedostojen koko on erittäin suuri (n. 1,5-2gb per kuva), joten ilman erittäin suuren kapasiteetin omaavan pilvipalvelun käyttöä kannettava kovalevy on ainoita järkeviä tapoja siirtää kuvia.

Kaikkiaan vajaa 80 näytettä sisältäneestä sarjasta valittiin testianalyysiin 10 näytettä satunnaisesti. Nämä näytteet analysoitiin alla olevan yleiseksikin analysointiohjeeksi kirjoitetun prosessin mukaisesti. Tämä prosessi on esitelty tarkemmin liitteenä 1 (suomeksi) ja liitteenä 2 (englanniksi).

Ylätasolla analysointivaiheet olivat seuraavat:

1. Perusta projekti analysoitavia näytekuvia varten ja lataa tarvittavat scriptit
2. Suorita “set staining vectors” scripti
3. Merkitse edustavat alueet kustakin näytteestä ja määritä näytesarjalle sopivat parametrit
4. Suorita “analysis” -scripti ja hienosäädä parametrejä, jos tarpeellista
5. Suorita “results to files” ja “summary table” -scriptit tulosten viemiseksi Exceliin
6. Nauti tuloksista

7 Tulokset

7.1 Vertailu manuaaliseen analyysiin

Näytteistä laskettiin ohjelmalla tietty tunnusluku, H-score, joka kuvastaa valitun näytteen kohdan värjäytymisen intensiteettiä. Tarkastelu tehtiin tässä testi-analyysissä tulehdussolujen osalta. H-scorea verrattiin sitten aiemmin kahden tutkijan tekemään manuaaliseen intensiteettipisteytykseen. Molemmat tulokset on esitetty taulukossa 1.

| Näytteen # | Värjäys | Skannauspäivä | QuPath H-Score | Tutkijoiden pisteytys |
|------------|---------|---------------|----------------|-----------------------|
| 1 | CHRNA7 | 5.9.2017 | 67,18 | 3 |
| 2 | CHRNA7 | 5.9.2017 | 114,74 | 3 |
| 3 | CHRNA7 | 5.9.2017 | 33,5 | 1 |
| 4 | CHRNA7 | 5.9.2017 | 45,72 | 1 |
| 5 | CHRNA7 | 5.9.2017 | 43,73 | 2 |
| 6 | CHRNA7 | 5.9.2017 | 93,76 | 2 |
| 7 | CHRNA7 | 5.9.2017 | 99,42 | 2 |
| 8 | CHRNA7 | 5.9.2017 | 115,29 | 3 |
| 9 | CHRNA7 | 5.9.2017 | 22,94 | 2 |
| 10 | CHRNA7 | 5.9.2017 | 35,55 | 1 |

| | |
|--------------|-------|
| Korrelaatio: | 0,697 |
|--------------|-------|

Taulukko 1. Ohjelmiston ja tutkijoiden pisteytyksen korrelaatio

Korrelaatio osoittautui merkittäväksi korrelaatiokertoimen ollessa 0,697, joten tietokoneavusteinen analyysi saatiin toimimaan oikeilla parametreilla hyvin.

7.2 Käyttäjien kokemukset

Tulokset valitun ohjelmiston ja analysointiohjeiden käytöstä ovat olleet rohkaisevia, joskin rajoitteitakin on havaittu. Ohjelman käyttöön perehtyi aluksi tämän tutkielman kirjoittaja, ja sen käytön perusteella luotiin ensimmäinen versio analysointiohjeista. Tässä vaiheessa ohjelmassa esiin nousi helppo valikko- ja symbolipohjainen käyttöliittymä sekä intuitiivinen työjonojen (scriptien) luontiominaisuus. Yhtenä selkeänä vahvuutena näyttäytyivät myös ohjelmiston käyttöohje- ja tukifoorumit, joista ongelmiin löysi ratkaisun, tai jos ei löytänyt, niin pyytäessä sen sai nopeasti.

Heikkoutena ensikäytössä näyttäytyi valittavien analysointiparametrien välillä liiankin suuri määrä sekä analysointitulosten heikko ulosvietävyys QuPathista esimerkiksi taulukkolaskentaohjelmaan.

Analysointi-ohjetta tarkennettiin yhden tutkijan (AA) testikäytön perusteella, ja tämän jälkeen ohjelmaa käyttivät edellä mainittu tutkija sekä kaksi muuta tutkijaa (MS, AH). Vahvuudeksi tässä testissä mainittiin selkeät ja johdonmukaiset ohjeet ohjelman käytölle. Ohjelmassa nähtiin rajoitteita, mutta se toimii hyvin H-scoren laskentaan. Ohjelma on helppokäyttöinen ja sillä analysointialueiden tarkka merkitseminen onnistuu hyvin. Puutteena ohjelmistossa nähtiin selkeän peruutus-toiminnallisuuden puute, eli jos esimerkiksi tekee virheellisen laajennuksen analysointialueeseen, niin koko alueen joutuu poistamaan ja aloittamaan merkinnän alusta. Vielä tätäkin suuremmaksi puutteeksi voidaan nostaa tarve manuaalisesti valita analysointialueet sen sijaan, että ohjelma analysoisi koko näytteen relevantit alueet. Tässä ongelmaksi muodostuu analysointialueiden valinta siten, että ne varmasti edustavat koko näytettä. Valinta vaatii siis sekä tarkkuutta että kokemusta.

Edellä mainitut analysointiohjeet laadittiin sekä englanniksi että suomeksi. Suomenkielinen analysointiohje on esitetty tämän tutkielman liitteenä 1 ja englanninkielinen liitteenä 2.

8 Pohdintaa

Kuten yleensäkin uusien toimintatapojen ja työkalujen kohdalla, alustavasta selvityksestä jää monia asioita vielä avoimeksi. Näitä voidaan löytää valitusta ohjelmistosta, ohjelmiston käyttöönotosta ja itse analysoinnin toteutuksesta.

Ohjelmiston valintaa varten tein varsin hyvän haastattelukierroksen ja kattavan kirjallisuuskatsauksen. Toisaalta täytyy muistaa, että kokemukset rajoittuivat nyt muutamiin yliopistoihin ja lääketieteen sovellutuksiin. Samaisia digitaalisen analysoinnin ohjelmistoja käytetään myös muilla tieteenaloilla (kuten vaikka metsätyyppejä satelliittikuvista tunnistamaan ja luokittelemaan), joten olisi hyödyllistä tutkia, että onko muilta aloilta saatu selkeän ylivertaisia kokemuksia jostain toisesta sovelluksesta. Sovelluksista täytyy muistaa, että nyt kartoitettiin ilmaiskäyttöön suunniteltuja ohjelmistoja, mutta jos jossain kohtaa tulevaisuutta

yliopisto tai muu taho osoittaa rahoitusta tämän alueen digitalisointiin, niin kaupallisiinkin ohjelmistoihin on hyvä tutustua syvällisemmin.

Nyt testikäyttöön valitusta QuPath -ohjelmistosta kokemusten perusteella voidaan monien hyvien puolien lisäksi nostaa ainakin yksi heikkous; vaikka se toimii solujen tai käyttäjän rajaamien alueiden tunnistukseen hyvin, niin sitä käytettäessä tutkijan fokus saattaa ajautua liian varhain pois laajemmasta yleiskuvasta. Peilaten aiempiin tutkimuksiin, niin esimerkiksi Salo ym. vuoden 2013 tutkimuksessa (31) tyypin I kollageenin roolia prognostisena merkkiaineena kielisyövässä ei ohjelmalla olisi ollut helppo havaita, sillä molekyylin suhteellisella esiintymisellä kielisyöpäkudosleikkeiden eri alueella, (pinta vs invaasioalue) oli ennusteen kannalta merkitystä, ei kollageenin kokonaismäärällä. Lisäksi Almahmoudi ym. havaitsivat vuonna 2018, että nimenomaan solunulkoiset, ei solulimassa olevat, interleukiini 17F-sytokiinit ennustivat kielisyövän käyttäytymistä (32). Tämänkään yhteyden huomaamiseen QuPath-ohjelmisto ei olisi riittänyt. Summaten voi sanoa, että ohjelmisto soveltuu nykyisellään paremmin pienien fokuksoidujen *mikroarray*-tyyppisten näytteiden analysointiin kuin koko laajan kokokudosleikkeen (engl. *whole slide image*) sisältäviin näytteisiin.

Myös ohjelmiston oikeaa käyttöä tulisi tukea ja seurata. Ohjeet toki voidaan jaella sähköpostilla eri tutkimusryhmille, mutta huomattavasti parempi ratkaisu olisi yhdistää jonkinlainen ohjelmiston ja analyysin esittelytilaisuus (joko kasvotusten tai vaihtoehtoisesti verkkopalaverina) kullekin ryhmälle. Ohjeiden jakelun ohessa olisi hyvä myös osoittaa taho, jolta saa tukea, jos ja kun eri näytesarjoilla uusia haasteita ilmaantuu. Nämä molemmat haasteet olisi ratkaistavissa nimeämällä ja kouluttamalla QuPathin käytön pääkäyttäjää. Tämä voisi esimerkiksi olla nimetty tutkija, joka osa-aikaisesti itse itsensä kouluttaisi hyvälle tasolle ohjelmiston käytössä ja voisi pienellä allokaatiolla pitää koulutuksia ja tarjota käytön tukea.

Analysointi nähtiin toimivaksi tähän tutkielmaan valikoituneella näytesarjalla ja halutuilla tarkastelukohteilla, mutta uusia haasteita varmasti esiintyy, kun ohjelmaa otetaan laajemmin käyttöön. Tämä ei tarkoita, että ohjelmistossa tai ohjeissa olisi suoranaista vikaa, vaan on muistettava, että tutkimusasetelmat ja -näytteet poikkeavat toisistaan, joten ohjelmiston konfigurointi ja parhaiden analysointityökalujen valinta kuhunkin tarpeeseen on väistämätöntä. Toki

analysointi onnistuu nyt luotujen ohjeiden avulla pienin muutoksen samanlaisiin sarjoihin, mutta tilanne on eri, jos halutaan analysoida täysin toisia indikaattoreita, kuten vaikkapa etäisyyksiä, pinta-aloja tai solujakautumisen aktiivisuutta.

Pitää vielä muistaa, että QuPath-ohjelmisto elää voimakasta kehitysvaihetta ja toisaalta sen kaikkiin nykyisiinkään hienouksiin ei tässä tutkimuksessa päästy perehtymään, joten sen kehityssivujen seuranta on lämpimästi suositeltua.

9 Kiitokset

Lopuksi haluan esittää lämpimät kiitokseni tässä tutkielmassa minua tukeneille henkilöille; tohtori Fang Zhaolle, LL Oscar Brückille, professori Markus Mäkiselle, dosentti Juha Väyryselle ja patologian erikoislääkäri Juha Näpänkankaalle haastatteluista ja keskusteluista. HLT Ahmed Al-Samadille, HLK Aini Hyytiäiselle ja HLK Meri Sieviläiselle ohjelmiston testikäytöstä ja arvokkaista kommentteista, ja erityisesti professori Tuula Salolle selkeästä, välittömästä ja huumorilla sävytetystä ohjaustyöstä.

Lähdeluettelo

(1) Rasband W. ImageJ. 2018; Available at: <https://imagej.nih.gov/ij/index.html>. Accessed 14.8., 2018.

(2) Zhao F. Haastattelu: Digitalisen analysoinnin käytön nykytila. 2017 20.6.2017.

(3) Väyrynen JP, Vornanen JO, Sajanti S, Böhm JP, Tuomisto A, Mäkinen MJ. An improved image analysis method for cell counting lends credibility to the prognostic significance of T cells in colorectal cancer. *Virchows Arch* 2012 May;460(5):455-465.

(4) Guirado R, Carceller H, Castillo-Gómez E, Castrón E, Nacher J. Automated analysis of images for molecular quantification in immunohistochemistry. *Heliyon* 2018 06/01;4(6):e00669.

(5) Bergmeister KD, Groger M, Aman M, Willensdorfer A, Manzano-Szalai K, Salminger S, et al. Automated muscle fiber type population analysis with ImageJ of whole rat muscles using rapid myosin heavy chain immunohistochemistry. *Muscle Nerve* 2016 Aug;54(2):292-299.

(6) Fulawka L, Halon A. Proliferation Index Evaluation in Breast Cancer Using ImageJ and ImmunoRatio Applications. *Anticancer Res* 2016 Aug;36(8):3965-3972.

(7) Grishagin IV. Automatic cell counting with ImageJ. *Anal Biochem* 2015 Mar 15;473:63-65.

(8) Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 2012 06/28;9:671.

(9) Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, et al. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* 2017 11/29;18(1):529.

(10) Wiesmann V, Franz D, Held C, Munzenmayer C, Palmisano R, Wittenberg T. Review of free software tools for image analysis of fluorescence cell micrographs. *J Microsc* 2015 Jan;257(1):39-53.

(11) Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nature Methods* 2012 06/28;9:676.

(12) de Chaumont F, Dallongeville S, Olivo-Marin J, editors. ICY: A new open-source community image processing software. 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro; 2011.

- (13) de Chaumont F, Dallongeville S. Icy: an open community platform for bioimage informatics. 2018; Available at: <http://icy.bioimageanalysis.org/>. Accessed 14.8., 2018.
- (14) Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. *Scientific Reports* 2017 12/04;7(1):16878.
- (15) Carpenter A, Jones T, Kametsky L, Goodman A, McQuin C, Et. al. CellProfiler | Free open-source software for measuring and analyzing. 2018; Available at: <http://cellprofiler.org/>. Accessed 14.8., 2018.
- (16) Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* 2006 10/31;7(10):R100.
- (17) McQuin C, Goodman A, Chernyshev V, Kametsky L, Cimini BA, Karhohs KW, et al. CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biology* 2018 07/03;16(7):e2005970.
- (18) Tollemar V, Tudzarovski N, Boberg E, Tornqvist Andren A, Al-Adili A, Le Blanc K, et al. Quantitative chromogenic immunohistochemical image analysis in cellprofiler software. *Cytometry A* 2018 Aug 8.
- (19) Zhang H, Ericsson M, Virtanen M, Westrom S, Wahlby C, Vahlquist A, et al. Quantitative image analysis of protein expression and colocalisation in skin sections. *Exp Dermatol* 2018 Feb;27(2):196-199.
- (20) BitPlane. Imaris - 3D and 4D Real-Time Interactive Microscopy Visualization Software | Imaris. 2018; Available at: <http://www.bitplane.com/Imaris/Imaris>. Accessed 14.8., 2018.
- (21) Choudhry P. High-Throughput Method for Automated Colony and Cell Counting by Digital Image Analysis Based on Edge Detection. *PLoS One* 2016 Feb 5;11(2):e0148469.
- (22) Plancoulaine B, Laurinaviciene A, Meskauskas R, Baltrusaityte I, Besusparis J, Herlin P, et al. Digital immunohistochemistry wizard: image analysis-assisted stereology tool to produce reference data set for calibration and quality control. *Diagn Pathol* 2014;9 Suppl 1:S8-1596-9-S1-S8. Epub 2014 Dec 19.
- (23) Loughrey MB, Bankhead P, Coleman HG, Hagan RS, Craig S, McCorry AMB, et al. Validation of the systematic scoring of immunohistochemically stained tumour tissue microarrays using QuPath digital image analysis. *Histopathology* 2018 Aug;73(2):327-338.
- (24) Tolonen T, Näpänkangas J, Isola J. Kliininen patologia virtuaalimikroskopian kynnyksellä. *Duodecim* 2015;131(21):1981-1987.

(25) Näpänkangas J, editor. Esitelmä: Kokemuksia Oulun diagnostisesta työstä ja patologian digitoinnista. ; 8.6.2017; : Näpänkangas, Juha; 2017.

(26) Näpänkangas J. Digitaalipatologia – mitä se on? Lääkärilehti 2018 2.2.2018;73(5/2018):249.

(27) Brück O. Haastattelu: Digitalisen analysoinnin käytön nykytila. 2017 22.6.2017.

(28) Väyrynen J, Mäkinen M. Haastattelu: Digitalisen analysoinnin käytön nykytila. 2017 22.8.2017.

(29) Näpänkangas J. Haastattelu: Digitalisen analysoinnin käytön nykytila. 2017 22.8.2017.

(30) Garvin S, editor. Seminaari: Kliinisen patologian nykynäkymiä / Case Linköping. ; 30.8.2017; : Garvin, Stina; 2017.

(31) Salo S, Bitu C, Merkkä K, Nyberg P, Bello IO, Vuoristo J, et al. Human Bone Marrow Mesenchymal Stem Cells Induce Collagen Production and Tongue Cancer Invasion. PLoS ONE 2013 09/02;8(10):e77692.

(32) Almahmoudi R, Salem A, Sieviläinen M, Sundquist E, Almangush A, Toppila-Salmi S, et al. Extracellular interleukin-17F has a protective effect in oral tongue squamous cell carcinoma. Head Neck 2018 May 14.

Liite 1: QuPath – Analysointiohje

1.1 Ohje ylätasolla

Nämä ohjeet on luotu skannattujen näytteiden värjäytyneiden solujen ja värjäytymisen intensiteetin analysointiin manuaalisesti valituilta näytealueilta. Edellä mainitut analysoitavat ominaisuudet yhdistetään yhdeksi indikaattoriksi – H-scoreksi. Nämä ohjeet pätevät käytettäessä avoimen lähdekoodin analysointiohjelmistoa nimeltä QuPath.

Ylätasolla analysointivaiheet ovat seuraavat:

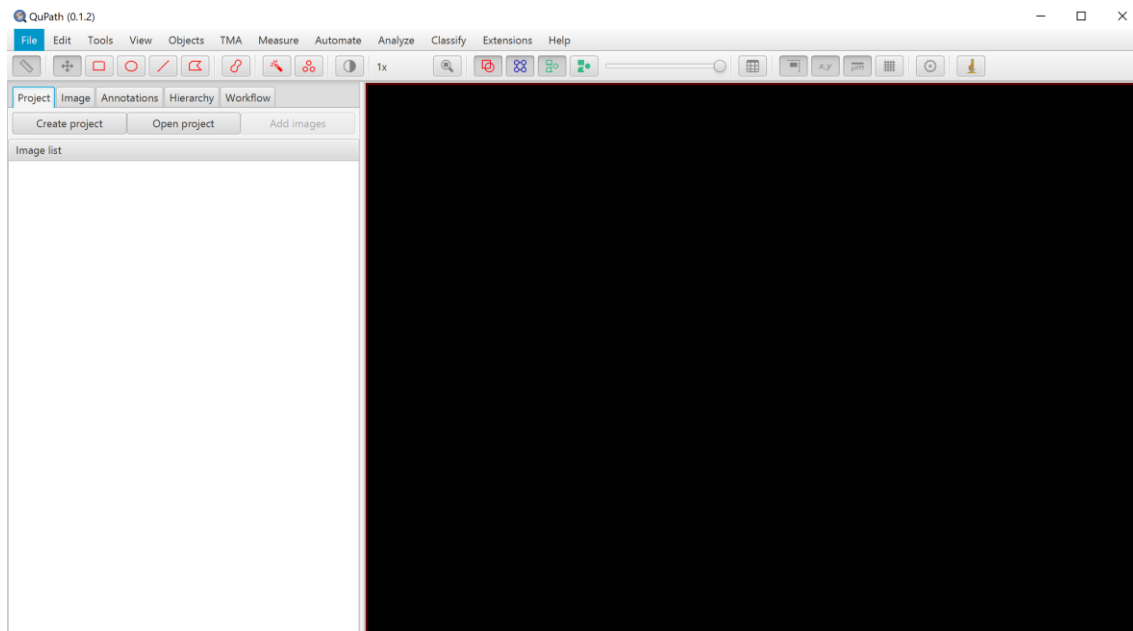
7. Perusta projekti analysoitavia näytekuvia varten ja lataa tarvittavat scriptit
8. Suorita “set staining vectors” scripti
9. Merkitse edustavat alueet kustakin näytteestä ja määritä näytesarjalle sopivat parametrit
10. Suorita “analysis” -scripti ja hienosäädä parametrejä, jos tarpeellista
11. Suorita “results to files” ja “summary table” -scriptit tulosten viemiseksi Exceliin
12. Nauti tuloksista

Kaikki nämä vaiheet käydään tarkemmin läpi alla.

1.2 Työvaiheet tarkemmin

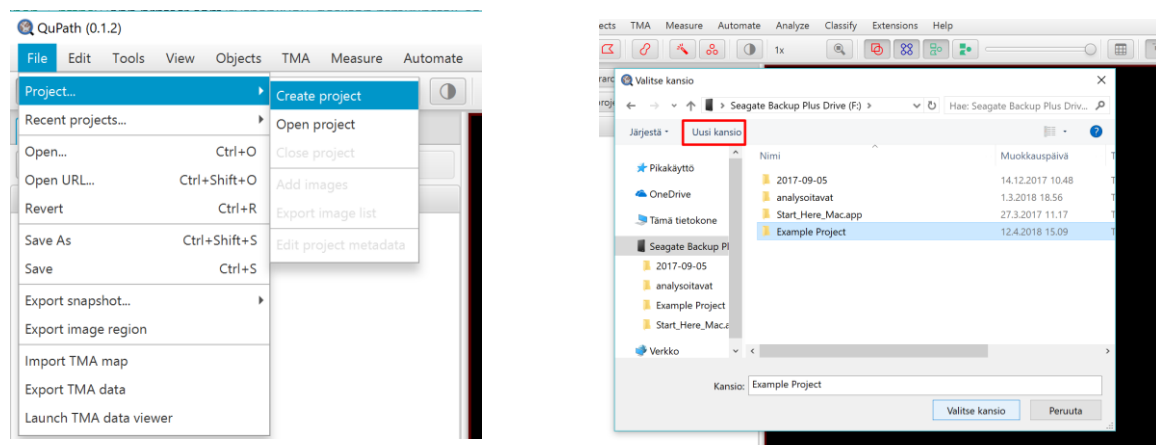
1. Perusta projekti analysoitavia näytekuvia varten ja lataa tarvittavat scriptit

Avaa aluksi QuPath. Perusnäkyvän pitäisi olla alla olevan kaltainen:



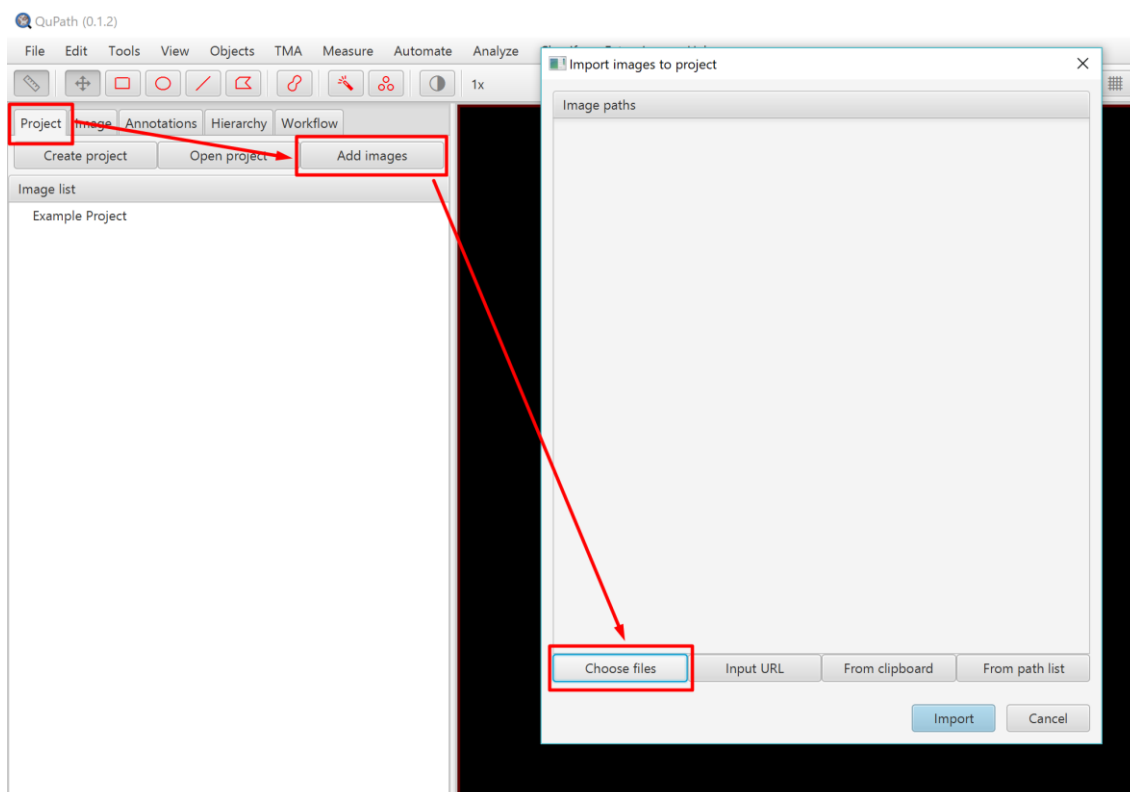
Kuva 1. QuPathin perusnäkö

Valitse “Create project” ja luo uusi kansio projektille. Kansio voi sijaita joko samalla laitteella (esim. kannettava kovalevy) kuin näytekuvat tai käytettävän tietokoneen kovalevyllä.



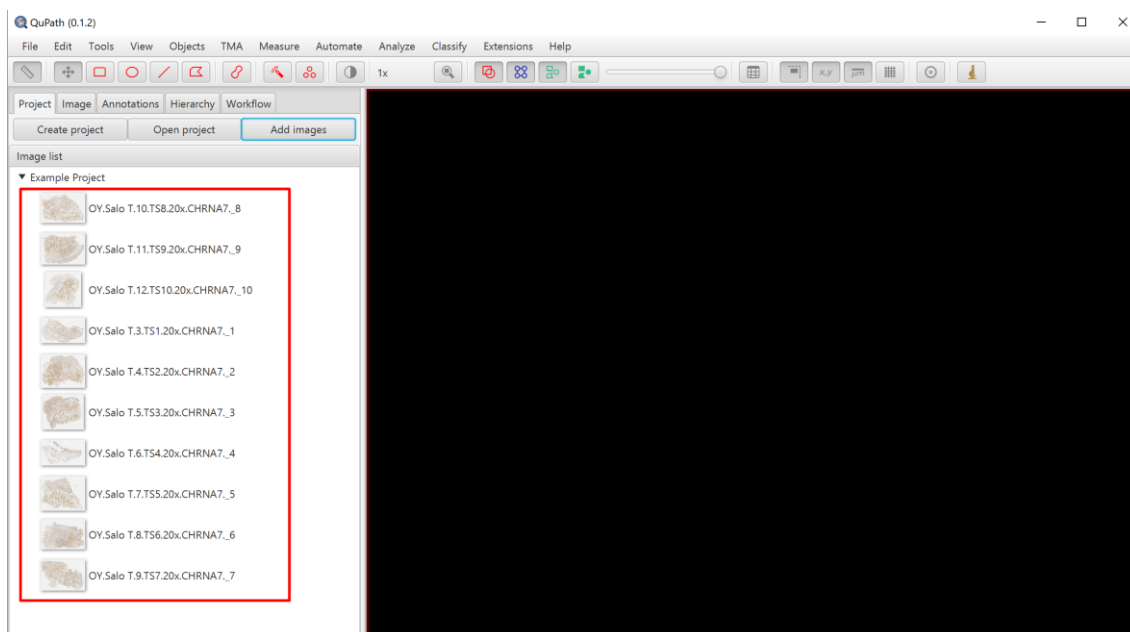
Kuva 2. Projektin ja projektikansion luominen

Seuraavaksi lisää haluamasi näytekuvat projektiin. Valitse “Project” -välilehti, sen jälkeen “Add Images” ja lopuksi valitse haluamasi lisäystapa. Helpoin tapa tähän on “Choose files”, jonka jälkeen käyttöjärjestelmän tyypillinen tiedostojen valintaikkuna aukeaa.



Kuva 3. Kuvien lisääminen projektiin

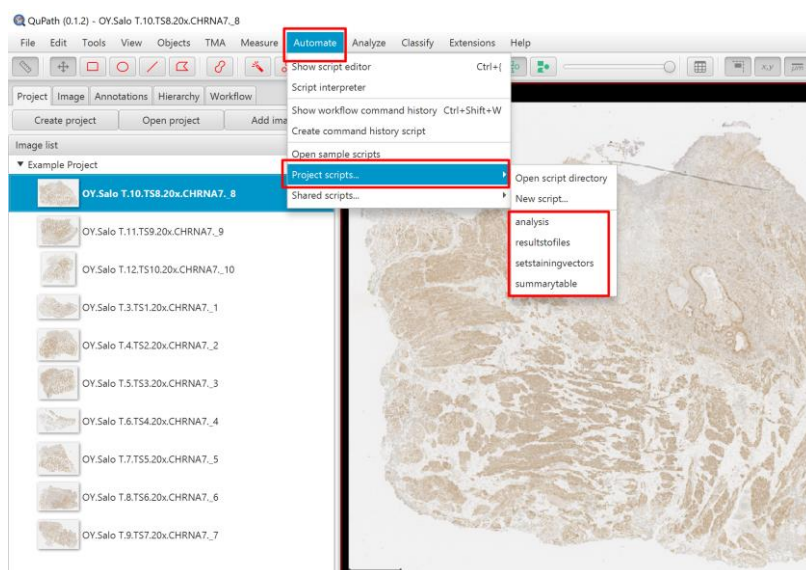
Tiedostojen valintaikkunassa tyypilliset monen tiedoston valintakomennot joko Ctrl- tai Shift-painike alas painettuna toimivat ja ovat näppäriä tähän. Valitse tiedostot ja paina "open". Tämän jälkeen QuPath näyttää latauksen tila -palkin, ja kun lataus on valmis, projektin kuvien tulisi näkyä kuten alla:



Kuva 4. Kuvien näkyminen projektissa lisäyksen jälkeen

Lisähuomautus: Nyt voi pikaisesti tarkastella näytteitä joko viemällä hiiren osoittimen leijumaan (hover) kuvan nimen päälle (pieni esinäkö) tai kaksoisklikkaamalla (kuva avautuu analysointi-ikkunaan).

Seuraavaksi kopioi analysoinnissa tarvittavat scriptit projektikansioon. Tämä vaihe suoritetaan QuPathin ulkopuolella käyttöjärjestelmän omassa tiedostonhallinta-näkymässä. Avaa kopioitavien scriptien lähdekansio (tässä tapauksessa: kannettava kovalevy → “H-score scripts”). Valitse kaikki scriptit ja kopioi (Ctrl-C). Mene projektikansioon, jonka yhtenä alikansiona näet automaattisesti luodun “scripts”-nimisen kansion. Mene tähän kansioon sisään ja sijoita (Ctrl-V) scriptit sinne. Voit tämän jälkeen varmistua scriptien latauksen olevan kunnossa QuPathissa viemällä hiiren leijumaan “Automate” → “Project scripts” ja täällä sinun tulisi nähdä scriptit listattuna ikkunassa. Tässä kohtaa ei ole syytä painaa mitään scripteistä.

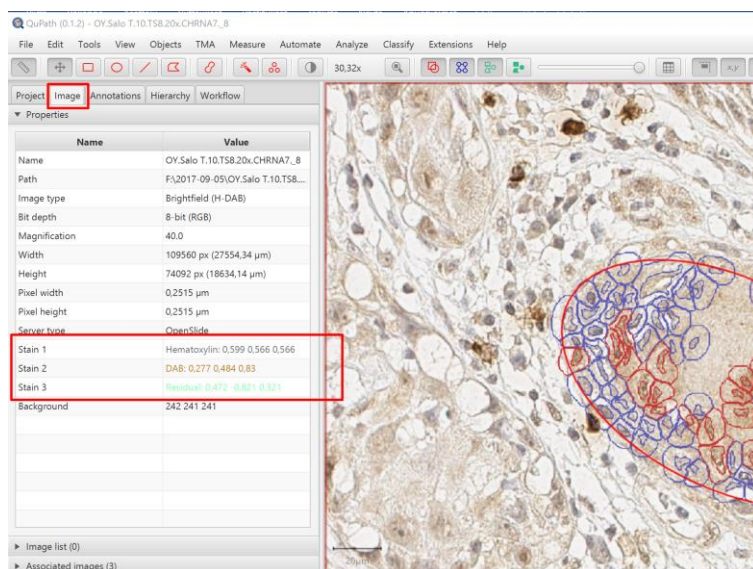


Kuva 5. Scriptien näkyminen QuPathissa latauksen jälkeen

2. Suorita “set staining vectors” scripti

Tällä työvaiheella ohjelmisto hienosäädetään erottelemaan värjäytynyt näytealue värjäytymättömästä ja kuvan taustasta. Toisin sanoen näytekuvasa on tyypillisesti 3 erillistä värjäytymisen “kerrosta”, jotka ohjelma erottaa toisistaan. Näiden kerrosten jako voidaan nähdä ja testata yksinkertaisesti painamalla numeronäppäimiä 1,2,3 ja 4 kun jokin näytekuva on auki analysointi-ikkunassa.

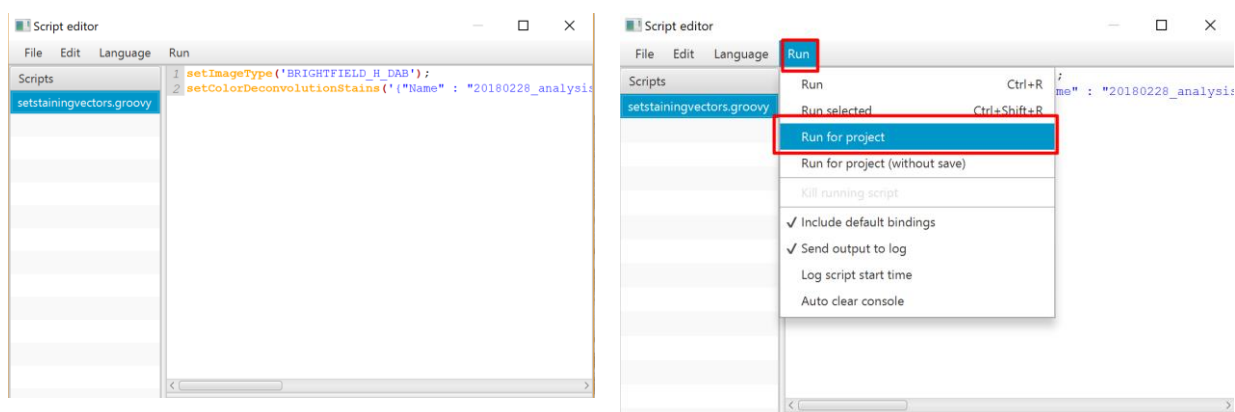
(1 = alkuperäinen kuva, 2 = värjäytymätyyppi 1 (tässä esimerkissä hematoksyliini), 3 = värjäytymätyyppi 2 (tässä esimerkissä DAB), 4 = värjäytymätyyppi 3 (tässä esimerkissä ja muutenkin tyypillisesti ”jäännös” eli tausta). Katso kuva alta:



Kuva 6. Värjäytymiskerrostien tutkiminen

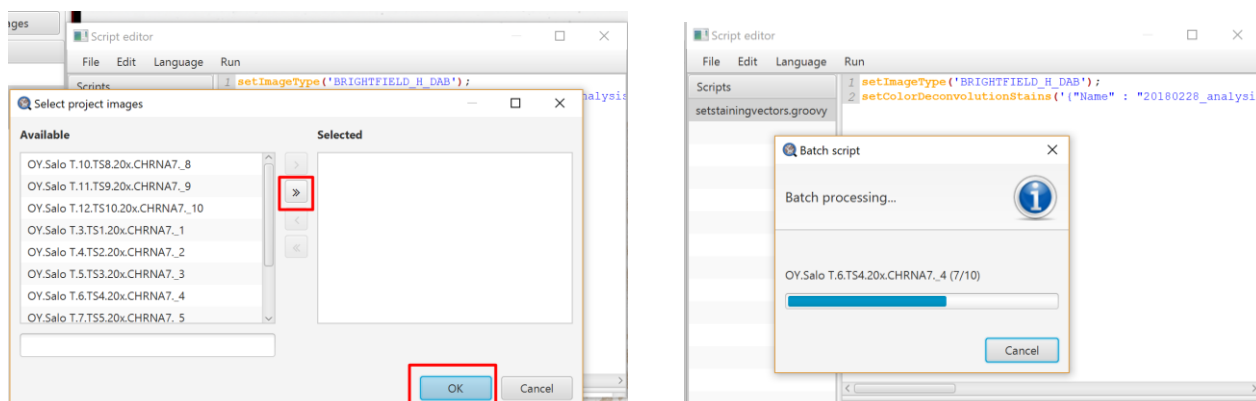
Värjäytymiskerrostien tunnistusvektoreiden (staining vectors) asetukseen on luotu yksinkertainen scripti ja se pitää suorittaa nyt. Tulee huomata, että kyseisen staining vectors -scriptin arvot toimivat optimaalisesti vain esimerkin näytekuvasarjalle, ja kun käytetään toista näytesarjaa, niin perusteellisempi vektoreiden asetus on tarpeen (tähän löytyy ohjeet QuPathin wiki-sivuilla: <https://github.com/qupath/qupath/wiki/Preprocessing>).

Valitse “Automate” → “Project scripts” → “setstainingvectors”. Tämän jälkeen alla näkyvä ikkuna aukeaa:



Kuva 7. Scripti-ikkuna ja scriptin suorittaminen projektille

Valitse sitten “Run” → “Run for project”. Seuraavanlainen dialogi-ikkuna aukeaa:

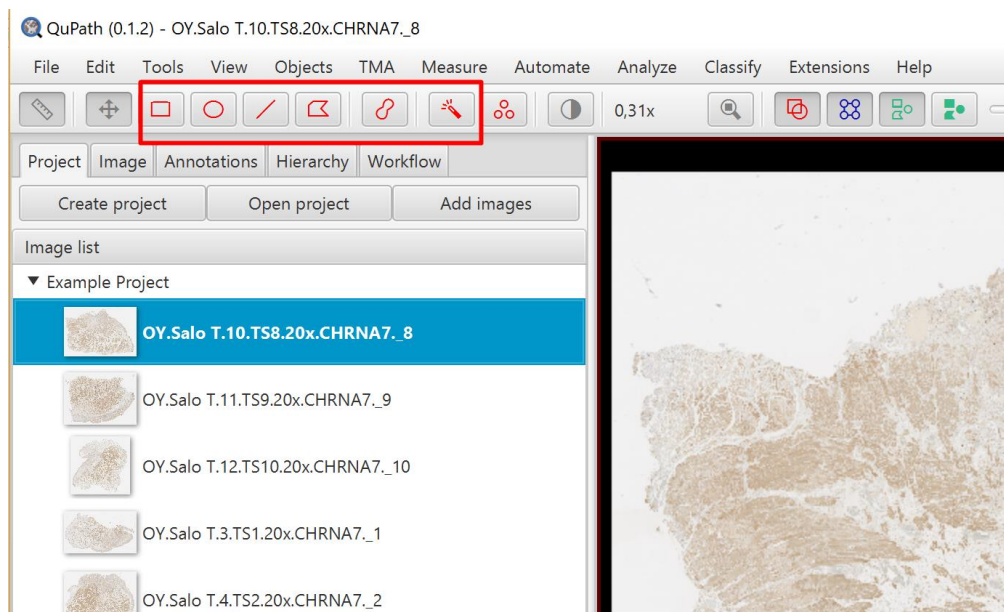


Kuva 8. Dialogi scriptin suorittamiseksi projektille

Tämä ikkuna pyytää valitsemaan kuvat projektista, joille valittu scripti ajetaan. Valitse kaikki tuplanuolella (tämä siirtää kaikki kuvien nimet oikean puoleiseen ikkunaan) ja klikkaa “ok”. Ohjelma näyttää edistymis-ikkunan kuten yllä olevassa kuvassa. Ajon loputtua script editor -ikkunan voi tältä erää sulkea.

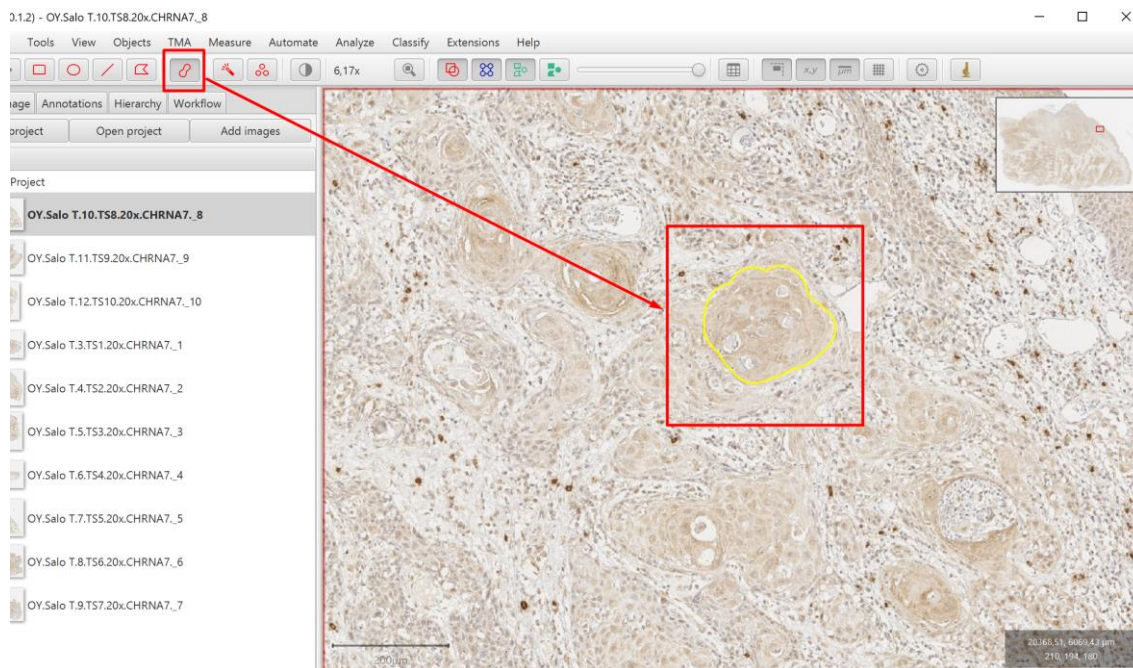
3. Merkitse edustavat alueet kustakin näytteestä ja määritä näytesarjalle sopivat parametrit

Seuraavaksi tulee valita analysoitavat alueet kustakin projektin kuvasta. Näitä alueita kutsutaan QuPathin terminologiassa englannin termillä *annotations*. Alueiden valinta voidaan tehdä erilaisilla työkaluilla, jotka on korostettu alla olevassa kuvassa:



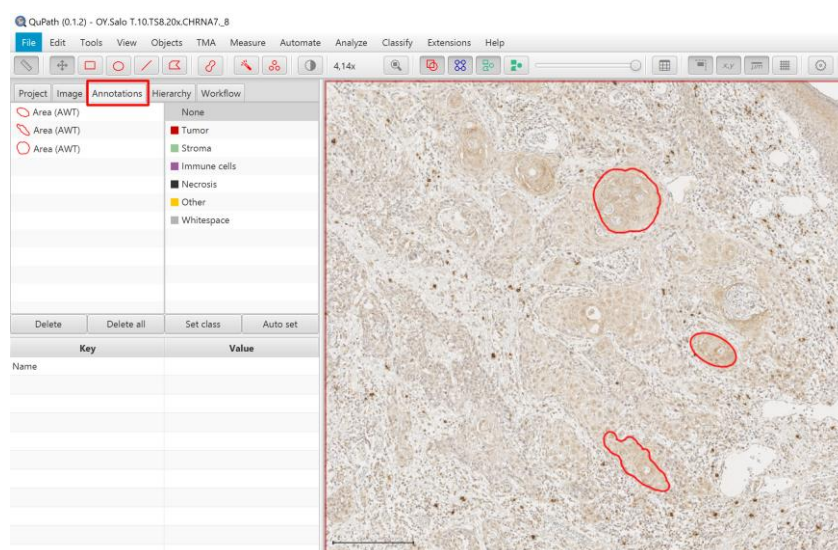
Kuva 9. Alueiden valintatyökalut

Tuplaklikkaa ensimmäistä kuvaa ja se aukeaa viereiseen analysointi-ikkunaan. Sitten zoomaamalla tarkoituksenmukaisesti etsi haluamasi alue näytteestä ja valitse se yhdellä valintatyökalulla (On järkevää testata valintatyökaluja etukäteen, jotta saa ajatuksen kunkin toimintaperiaatteesta). Esimerkki valinnasta alla:



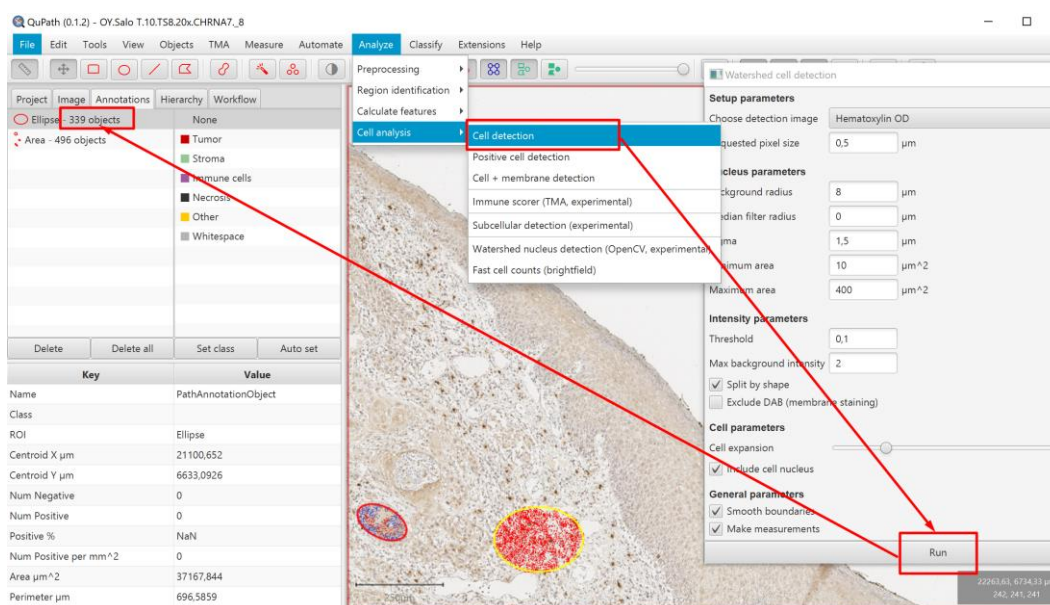
Kuva 10. Esimerkki alueen (annotation) valinnasta

Valitse yhteensä 3 erillistä näytealuetta (tämä määrä voi vaihdella riippuen näytesarjasta ja sovitusta analysointistrategiasta). Valinnat voidaan nähdä “Annotations” –välilehdeltä, ja yhden kuvan osalta ne voisivat näyttää esimerkiksi tältä:



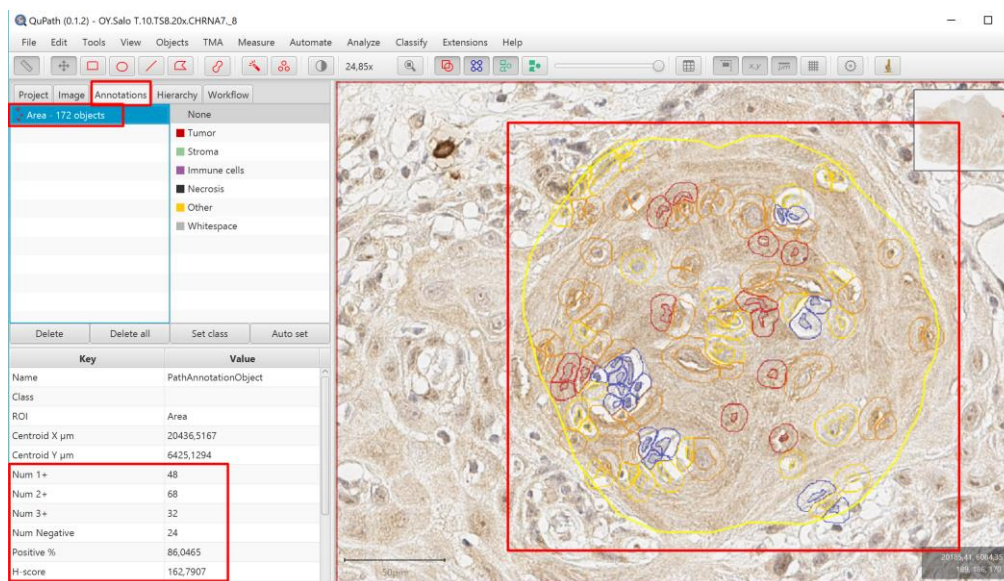
Kuva 11. Alueiden (annotations) näkyminen omalla välilehdellään

Riippuen analysointistrategiasta voi olla, että halutaan valita alueet, jotka pitävät sisällään karkeasti tietyn määrän soluja. Solujen määrän selvittämiseksi tietyn alueen (annotation) sisällä on olemassa nopea metodi: Varmista, että haluttu alue on aktiivinen, eli reunaväritään keltainen. Valitse sitten cell analysis → cell detection → “Run”. Ohjelma tunnistaa ja laskee solut ja näyttää tuloksen valitun alueen nimen vieressä (vasen puoli alla olevassa ruutukaappauksessa). Jos tulos on liian pieni tai liian suuri, niin poista alue ja tee uusi valinta tai muokkaa alkuperäistä aluetta.



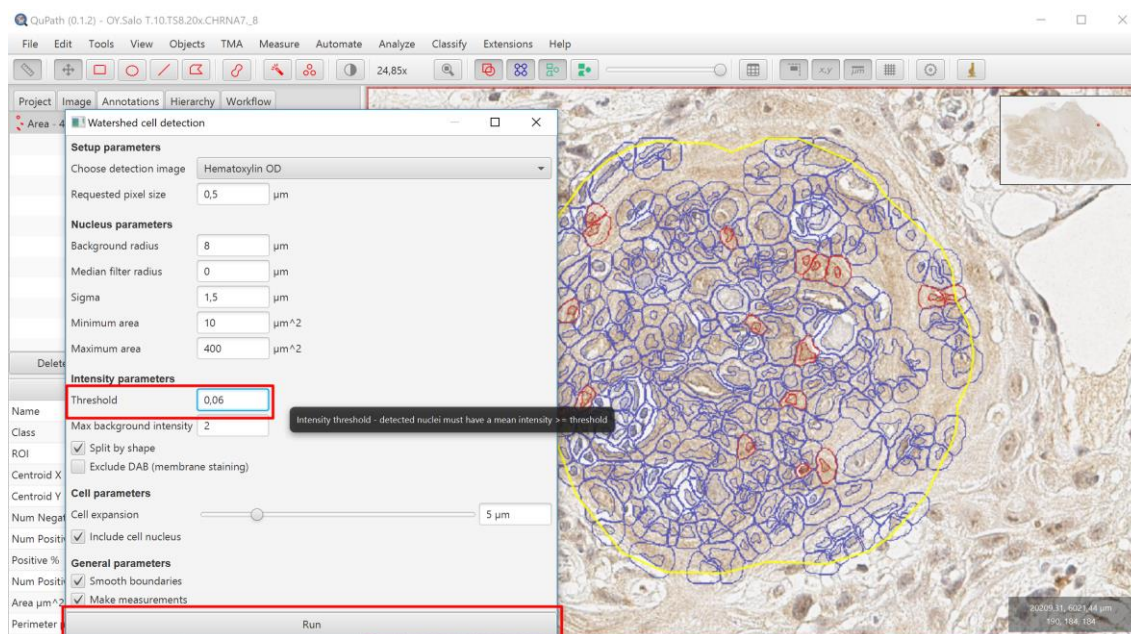
Kuva 12. Solumäärän selvittäminen alueesta

Joskus voi olla, että ohjelma ei tunnista soluja tarkoituksenmukaisesti kuten voidaan nähdä seuraavasta esimerkistä:



Kuva 13. Vajavainen solujen tunnistaminen

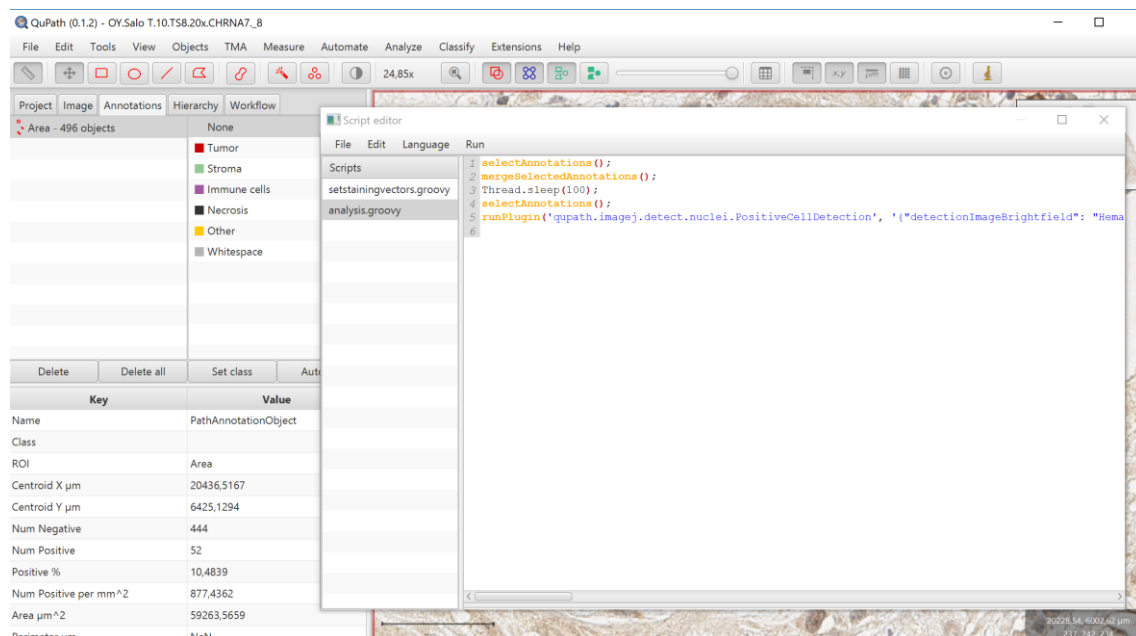
Tässä tapauksessa näyttää, että tunnistus ei ole toiminut ja iso osa soluista on jäänyt tunnistamatta. Prosessia voi hienosäätää ajamalla manuaalisesti yksinkertaista tunnistusta, joka esiteltiin ylempänä, ja tutkien, että miten tunnistus reagoi parametrien muutoksiin. Valitse jälleen “Analyze” → “Cell analysis” → “Cell detection” ja seuraava dialogi-ikkuna aukeaa.



Kuva 14. Soluntunnistuksen parametrien hienosäätäminen

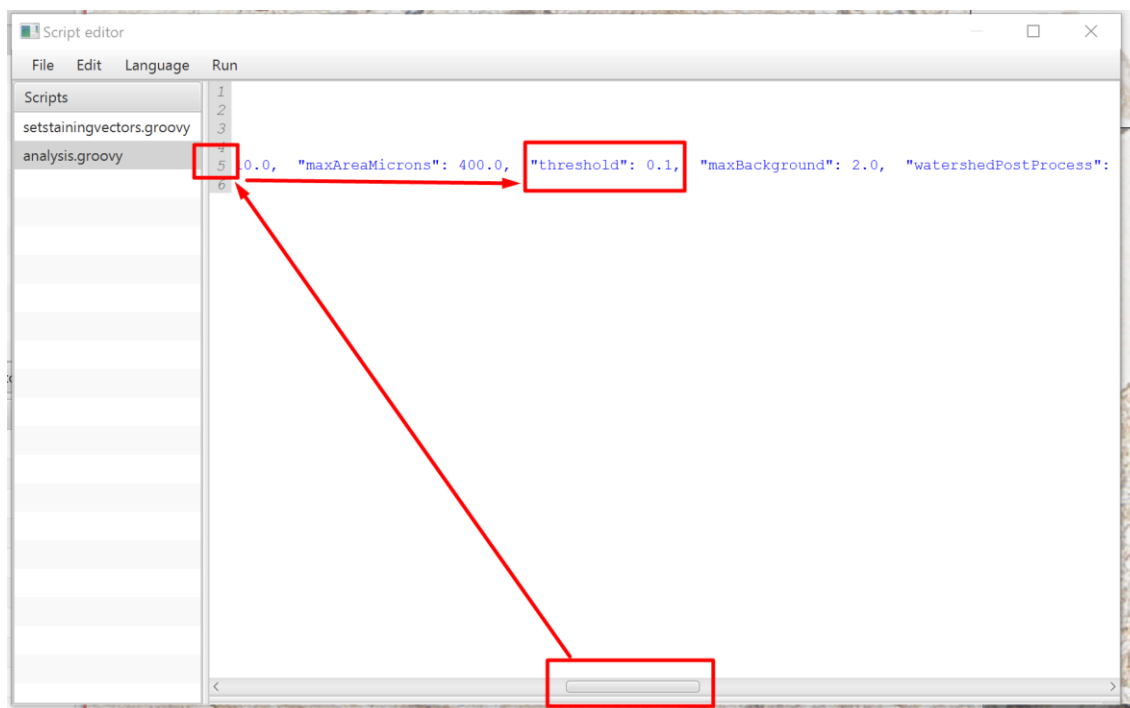
Useita parametrejä voi muokata, mutta tähän tarkoitukseen **intensiteetin** kynnysarvon (intensity threshold) hienosäätö yleensä auttaa. Oletusarvo tälle on aina 0,1. Säädä arvo esimerkiksi 0,06 ja paina "Run". Voit nähdä kuinka tunnistustulos muuttuu kuvassa (tulokset näytetty yllä olevan kuvan

oikeanpuoleisessa osassa). Kun olet tyytyväinen intensiteetin kynnysarvoon, tämä pitää päivittää projektin scripttiin, jotta samaa arvoa käytetään läpi kaikkien projektin näytekuviin. Mene projektin scriptteihin ja valitse “analysis” –scriptti. Alla oleva editori-näkymä aukeaa:



Kuva 15. Scriptien muokkauksen editori-näkymä

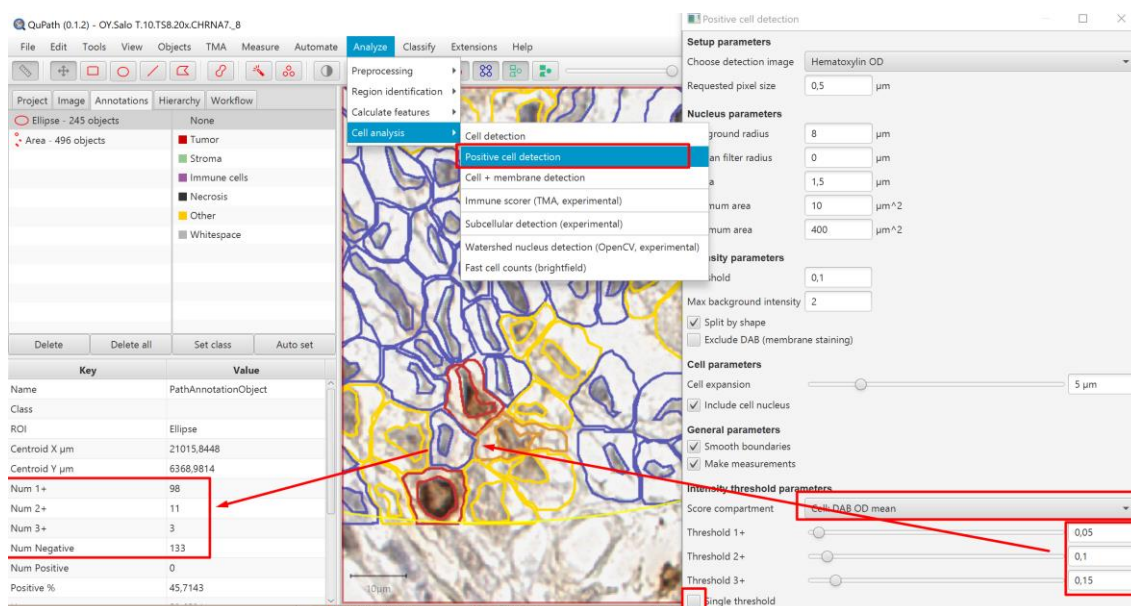
Tässä näkymässä vieritä sivusuunnan palkkia oikealla kunnes löydät riviltä 5 kynnysarvon parametrin. Päivitä tähän toimivaksi havaitsemasi arvo kirjoittamalla se näppäimistöllä.



Kuva 16. Kynnysarvon löytäminen ja päivittäminen scriptiin

Tämän jälkeen valitse scripti-editorissa “File” → “Save”. Tämä komento tallettaa uuden arvon scriptiin.

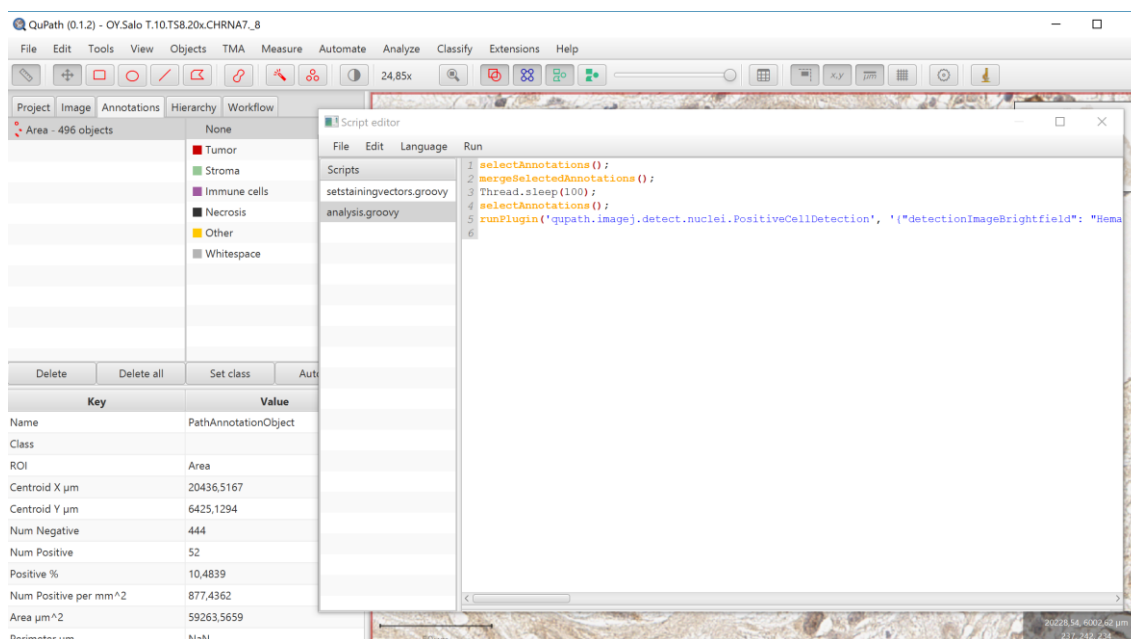
Muut seikat, jotka voivat tyypillisesti vaatia hienosäätöä ovat intensiteetin luokkien kynnysrajat (intensity threshold parameters). Näillä tarkoitetaan värjäytymisen intensiteettitasoja, jotka jakavat solut intensiteettiluokkiin negatiivinen, luokka 1 (heikko), luokka 2 (keskiverto) ja luokka 3 (vahva). Valitse jälleen haluamasi alue (annotation) ja valitse “cell analysis” → “**positive** cell detection”. Tästä dialogi-ikkunasta tiettyjä asioita tulee nyt muuttaa: Tuloksen laskualue (Score compartment) tulee valita vastaamaan analysointistrategiaa ja tässä esimerkissä se asetettiin kattamaan koko solun ja käyttämään intensiteetin keskiarvoa. Tämän jälkeen pitää poistaa ruksi “Single threshold” -valinnasta ja viimeiseksi säätää itse kynnysarvoja. Paina “Run” ja solut värikoodautuvat intensiteettiluokkansa mukaisesti: Sininen = negatiivinen, Keltainen = luokka 1 (heikko), Oranssi = luokka 2 (keskiverto), Punainen = luokka 3 (vahva). Solumäärät per luokka voidaan myös nähdä vasemmasta paneelistä.



Kuva 17. Parametrien säätö intensiteettiluokille

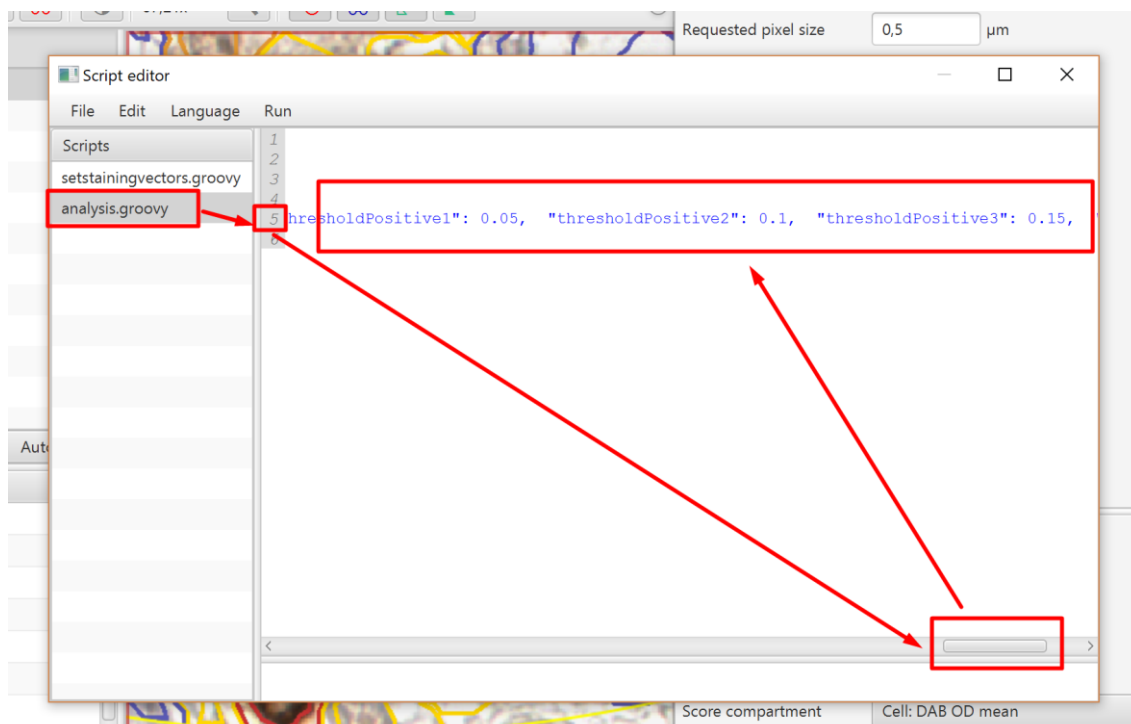
Miksi tämä säätö? Voit huomata silmämääräisesti tarkastellen ja perustuen kokemukseesi, että kynnsarvot tarvitsevat muutoksia ja ne pitää tietysti päivittää koskemaan kaikkia projektin kuvia. Kannattaa testata arvot sopiviksi yhdellä alueella (annotation) ja kirjata ne vaikkapa muistilapulle.

Kun olet tyytyväinen intensiteetin kynnsarvoihin nämä tulee päivittää projektin vastaavaan scriptiin. Mene jälleen “project scripts” ja valitse “analysis” –scripti. Vastaava editori-näkymä aukeaa:



Kuva 18. Script-editorin näkymä kynnsarvojen päivittämiseksi

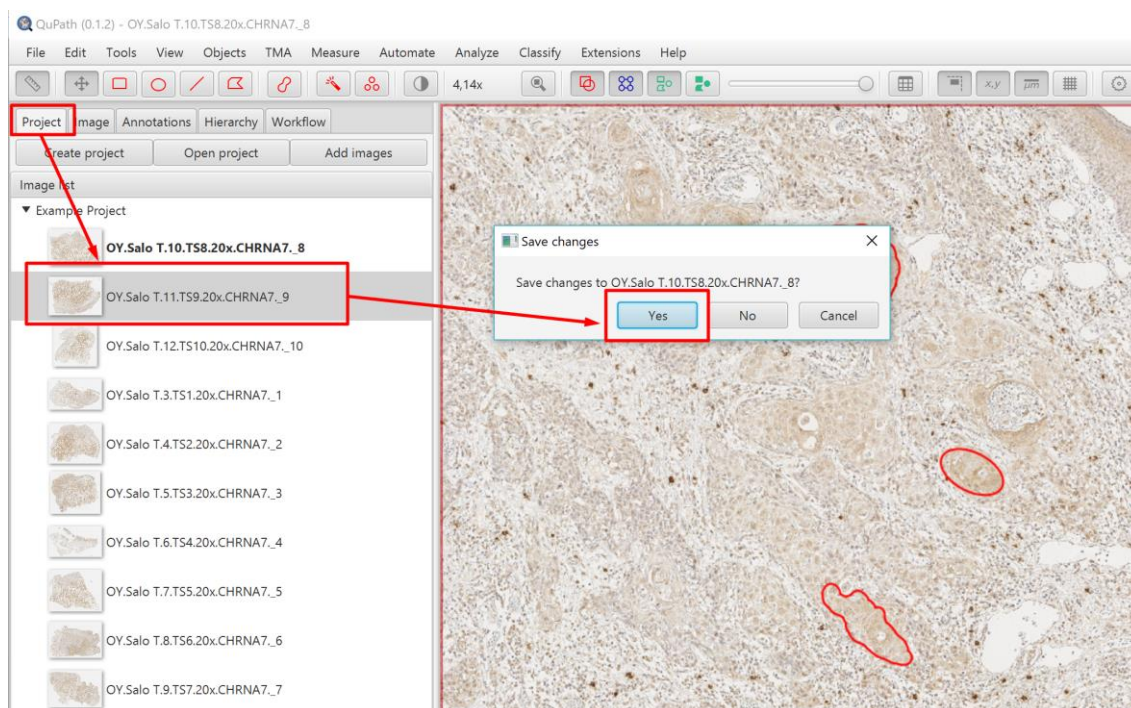
Vieritä näkymää oikealle kunnes löydät riviltä 5 kynnysarvoparametrit. Päivitä näihin toimivaksi havaitsemasi numeroarvot.



Kuva 19. Luokkien kynnysarvojen päivittäminen analysis-scriptiin

Tämän jälkeen valitse scripti-editorissa “File” → “Save”. Tämä tallettaa uudet arvot scriptiin. Scripti-editori voidaan nyt sulkea ja voit jatkaa aluevalintojen (annotations) tekemistä kyseiselle näytekuvalle.

Kun olet tyytyväinen ensimmäisen kuvan analyysialuevalintoihin, mene takaisin ”Project”-välilehdelle ja valitse seuraava kuva tuplaklikkaamalla. Ohjelma kysyy muutosten tallentamisesta ja tähän vastaa “Yes”.



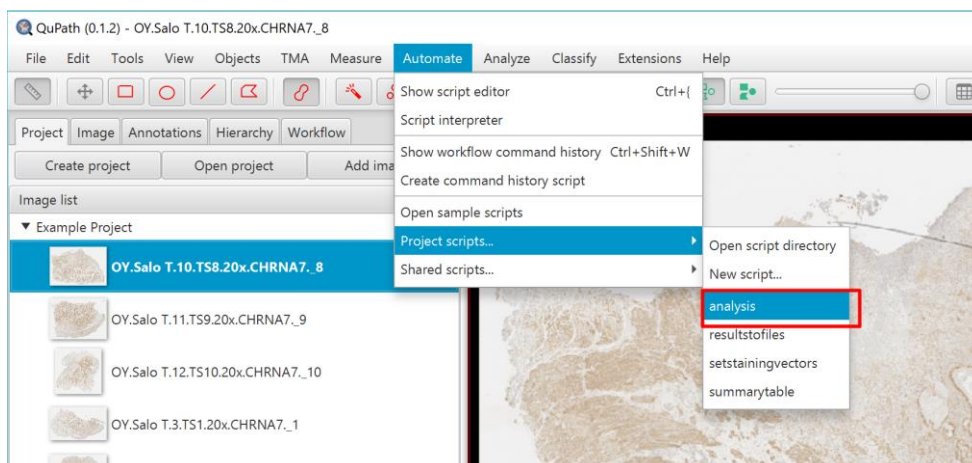
Kuva 20. Näytekuivissa eteneminen ja muutosten tallennus

Toista aluevalintojen tekeminen kaikille näytekuville projektissa. Huomioi, että tämä aluevalintojen tekeminen on kaikkein työläin vaihe koko prosessissa.

Lisähuomio: Jos teet virheen alueen valinnassa, niin valitse kyseinen virheellinen alue (se muuttuu keltaiseksi) ja valitse “Objects” → “Delete” → “Delete selected objects”. Valitettavasti QuPathissa ei ole suoranaista undo-toiminnallisuutta.

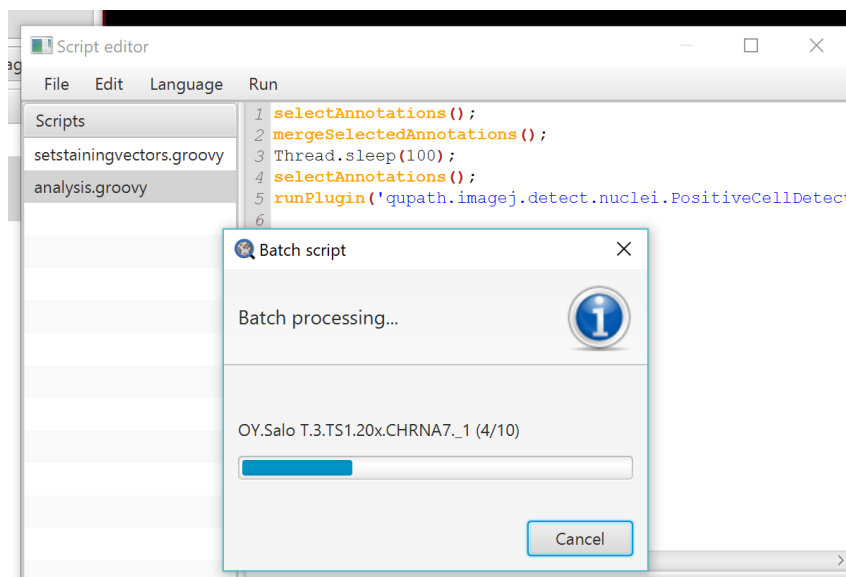
4. Suorita “analysis” -scripti ja hienosäädä parametrejä, jos tarpeellista

Tähän mennessä kaikki näytekuvat on siis saatu tallennettua kolmella näytealuevalinnalla. Palaa takaisin ensimmäiseen kuvaan ja valitse “Automate” → “Project scripts” → “analysis”. Tämä scripti yhdistää aluksi kunkin kuvan 3 aluetta yhdeksi ja tämän jälkeen tunnistaa ja luokittelee solut. Valitse jälleen “Run” → “Run for project”.



Kuva 21. Analysis-scriptin valinta

Scriptin käynnistyttyä alla olevan kaltainen edistymisikkuna näytetään:



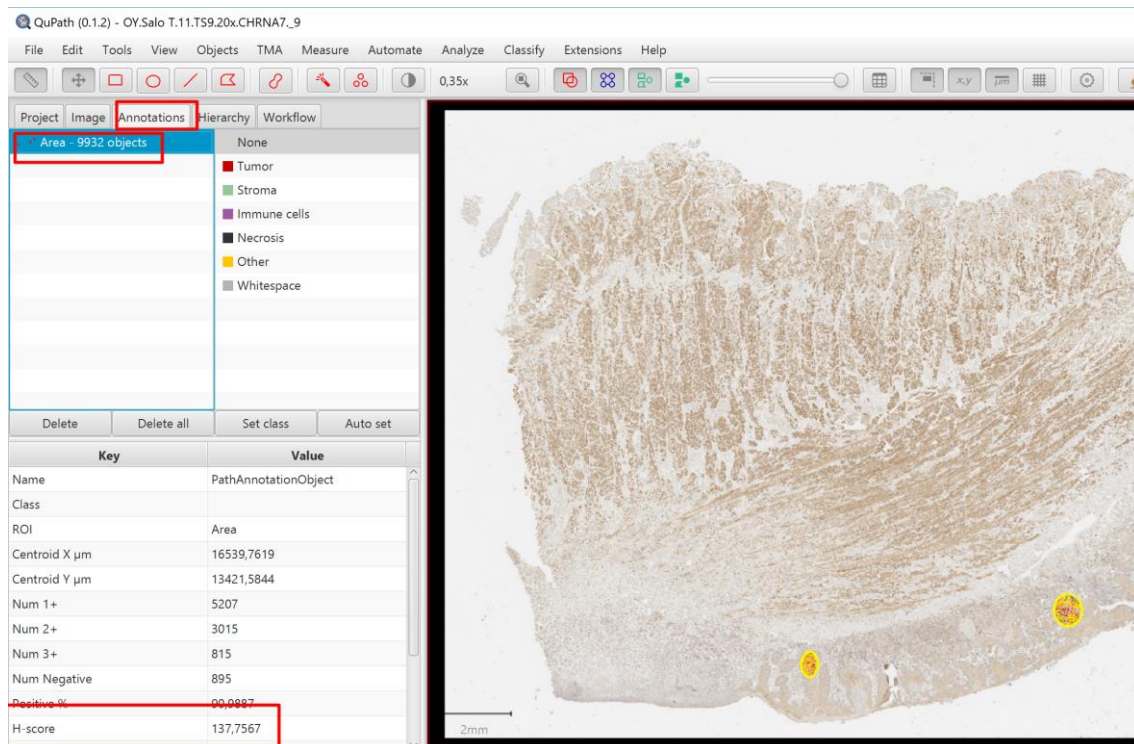
Kuva 22. Edistymisikkuna, kun scripti ajetaan koko projektille

Ajon päätyttyä scripti-editori voidaan sulkea. Tuloksia voidaan nyt tarkastella tuplaklikkaamalla haluttua kuvaa ja menemällä "Annotations" –välilehdelle ja tämän lisäksi tietysti zoomaamalla alueita itse analysointi-ikkunassa.

Entä jos jokin parametri vielä tarvitsisi muutosta? Aluksi on viisasta testata muutosten vaikutusta yhdellä alueella (annotation) manuaalisesti, kuten on kuvattu tämän ohjeen osiossa 3. Sitten kun uudet arvot ovat löytyneet, niin päivitä nämä analysis-scriptiin. Tämän jälkeen valitse taas "Run" and "Run for project". Tämä uudelleenajaa analyysin kaikille kuvilla päivitettyillä parametreillä. Voit jälleen tarkastella uusia tuloksia eri kuvista.

5. Suorita “results to files” ja “summary table” -scriptit tulosten viemiseksi Exceliin

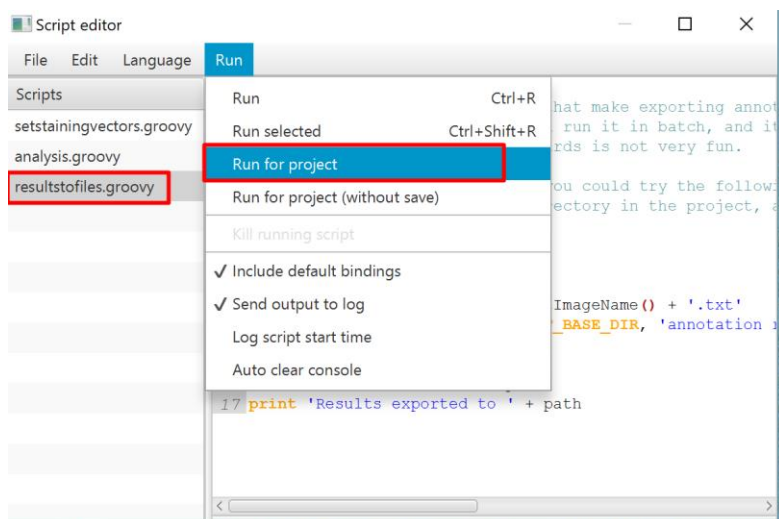
Kuten aiempien prosessivaiheiden jälkeen voidaan nähdä, H-score on jo valmiiksi laskettu kullekin kuvalle. Voit nähdä tämän menemällä “Annotations”-välilehdelle ja valitsemalla ainoan (yhdistetyn) alueen:



Kuva 23. H-scoren näkyminen valitulle alueelle

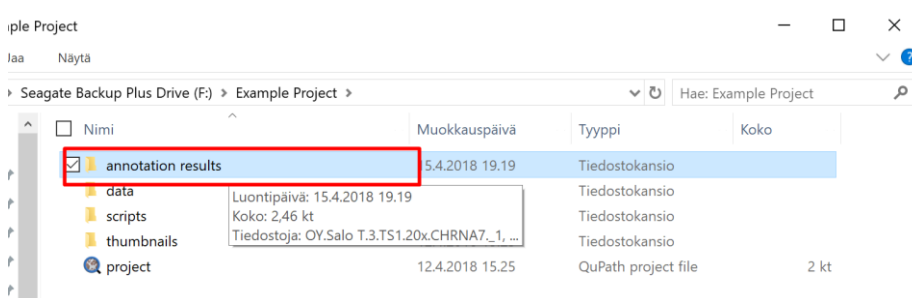
Tämän prosessivaiheen tarkoituksena on kerätä nämä analysointitulokset yhteen txt-tiedostoon. QuPath ei tällä hetkellä sisällä kunnollisia export-toiminnallisuuksia, joten tarvitaan 2 erillistä scriptiä.

Aja ensiksi “resultstofile”-scripti projektille:



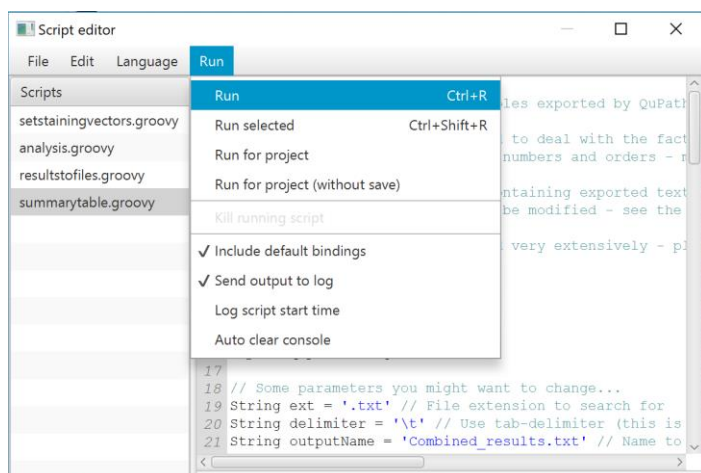
Kuva 24. "resultstofile"-scriptin ajaminen projektille

Tämä luo "annotation results" -alikansion aiemmin luomasi projektikansion alle:



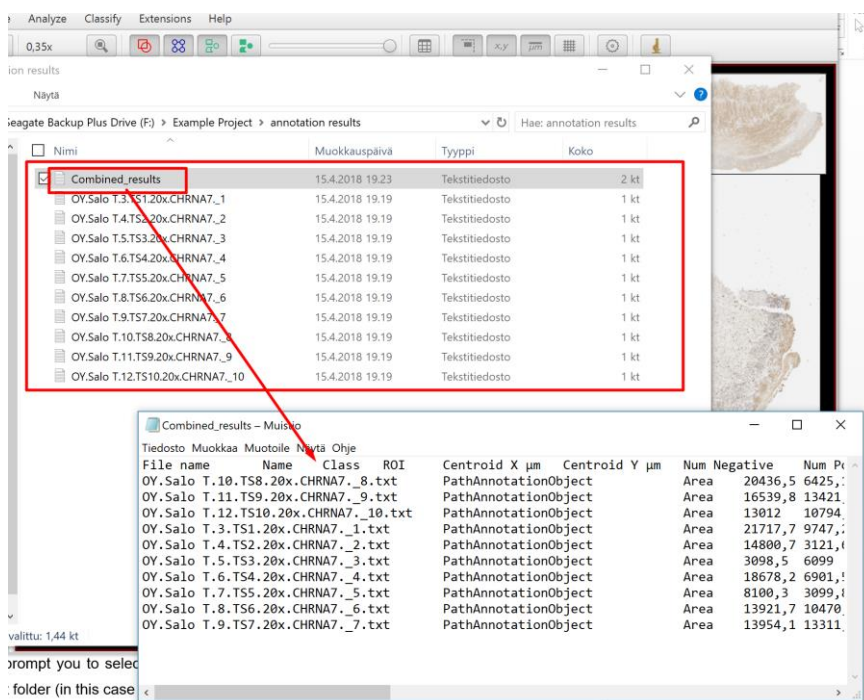
Kuva 25. Uusi alikansio projektikansiossa

Alikansiossa on nyt erilliset txt-tiedostot kunkin näytekuvan osalta. Seuraavaksi tulee ajaa "summarytable"-scripti. HUOMAA, että tämä tulee ajaa ainoastaan yksittäisellä "Run" -komennolla, ja ei ole väliä, että mikä projektin näytekuvista on auki taustalla. Tämä scripti yhdistää aiemmat tiedostot yhdeksi koontitiedostoksi.



Kuva 26. "Run"-komento "summarytable"-scriptille

Ohjelma pyytää valitsemaan kansion, jossa erilliset tulostiedostot sijaitsevat. Valitse oikea kansio (tässä tapauksessa “annotation results”). Tämän jälkeen voit mennä jälleen käyttöjärjestelmän tiedostonhallinnan puolella ja katsoa, että miltä kansion sisältö näyttää. Täällä tulisi olla nyt “combined results” –tekstitiedosto ja kun klikkaat sen auki eri näytekuviin tuloksien tulisi olla taulukkomuodossa.



Kuva 27. Koontitiedoston tulosten tarkastelu

Tämä tiedosto on luettavampi ja jatkoanalysoitavampi avattaessa se taulukkolaskentaohjelmassa, kuten MS Excelissä. Tässä esimerkiksi H-scoret ovat näitesti tutkittavissa tai poimittavissa:

| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |
|----|-------------------------------------|----------------------|-------|---------|------------|------------|-----------|------------|------------|----------|-----------|--------|--------|--------|---------|-------------|--------------|--------------|
| 1 | File name | Name | Class | ROI | Centroid X | Centroid Y | Num Negat | Num Positi | Positive % | Area µm² | Perimeter | Num 1+ | Num 2+ | Num 3+ | H-score | Allred prop | Allred inter | Allred score |
| 2 | OY.Salo T.10.TS8.20x.CHRNA7_8.txt | PathAnnotationObject | Area | 20436,5 | 6425,1 | 444 | 52 | 10,48 | 877,44 | 59263,6 | NaN | 5207 | 3015 | 811 | 137,76 | 5 | 1 | 6 |
| 3 | OY.Salo T.11.TS9.20x.CHRNA7_9.txt | PathAnnotationObject | Area | 16539,8 | 13421,6 | 895 | 90,99 | 90,99 | | 1267382 | NaN | | | | | | | |
| 4 | OY.Salo T.12.TS10.20x.CHRNA7_10.txt | PathAnnotationObject | Area | 13012 | 10794,2 | 1791 | 60,38 | 96,69 | | 825225,9 | NaN | 1353 | 752 | 625 | 104,67 | 4 | 2 | 6 |
| 5 | OY.Salo T.3.TS1.20x.CHRNA7_1.txt | PathAnnotationObject | Area | 21717,7 | 9747,2 | 188 | 82,08 | 96,69 | | 535386 | NaN | 1722 | 2134 | 164 | 191,96 | 5 | 2 | 7 |
| 6 | OY.Salo T.4.TS2.20x.CHRNA7_2.txt | PathAnnotationObject | Area | 14800,7 | 3121,6 | 556 | 82,08 | 96,69 | | 647320 | NaN | 933 | 814 | 795 | 159,83 | 5 | 2 | 7 |
| 7 | OY.Salo T.5.TS3.20x.CHRNA7_3.txt | PathAnnotationObject | Area | 3098,5 | 6099 | 3 | 99,86 | 99,86 | | 195315,1 | NaN | 66 | 204 | 181 | 283,46 | 5 | 3 | 8 |
| 8 | OY.Salo T.6.TS4.20x.CHRNA7_4.txt | PathAnnotationObject | Area | 18678,2 | 6901,5 | 1 | 99,95 | 99,95 | | 221356,1 | NaN | 140 | 553 | 121 | 256,23 | 5 | 3 | 8 |
| 9 | OY.Salo T.7.TS5.20x.CHRNA7_5.txt | PathAnnotationObject | Area | 8100,3 | 3099,8 | 233 | 78,89 | 78,89 | | 149145,3 | NaN | 367 | 380 | 12 | 135,78 | 5 | 2 | 7 |
| 10 | OY.Salo T.8.TS6.20x.CHRNA7_6.txt | PathAnnotationObject | Area | 13921,7 | 10470,8 | 18 | 99,7 | 99,7 | | 529944,9 | NaN | 257 | 1704 | 396 | 261,76 | 5 | 3 | 8 |
| 11 | OY.Salo T.9.TS7.20x.CHRNA7_7.txt | PathAnnotationObject | Area | 13954,1 | 13311,2 | 12 | 97,55 | 97,55 | | 85112 | NaN | 51 | 86 | 34 | 254,29 | 5 | 3 | 8 |

Kuva 28. H-scorejen tarkastelu MS Excelissä

6. Nauti tuloksista

Siinä kaikki. Jos joitain muutostarpeita herää, niin voit palata haluamaasi vaiheeseen ja tehdä muutokset esimerkiksi soluntunnistuksen parametreihin tai aluevalintoihin.

Loppuhuomaus: Nämä ohjeet ovat tyyppiä “vain seuraa ja usko”, mutta toki parempi lähestyminen olisi ymmärtää mitä ohjelma ja scriptit tekevät kussakin vaiheessa. Ohjelman käyttö on tehty erittäin helpoksi, joten perusteiden ja muutamien edistyneempien toimintojen opettelu ei vaadi liikaa aikaa. Perehtyäkseen tarkemmin QuPathiin on vahvasti suositeltavaa lukea ohjelman wiki-sivua: <https://github.com/qupath/qupath/wiki>

Liite 2: QuPath – Analysis Instructions (engl.)

2.1 Overview of the procedure

These instructions are created for analysing stained cells and the intensity of the staining from manually selected areas of the scanned samples. Ultimately these are combined as one indicator – H-score. The instructions apply for analysis done in open-source software called QuPath.

The procedure goes as follows on high level:

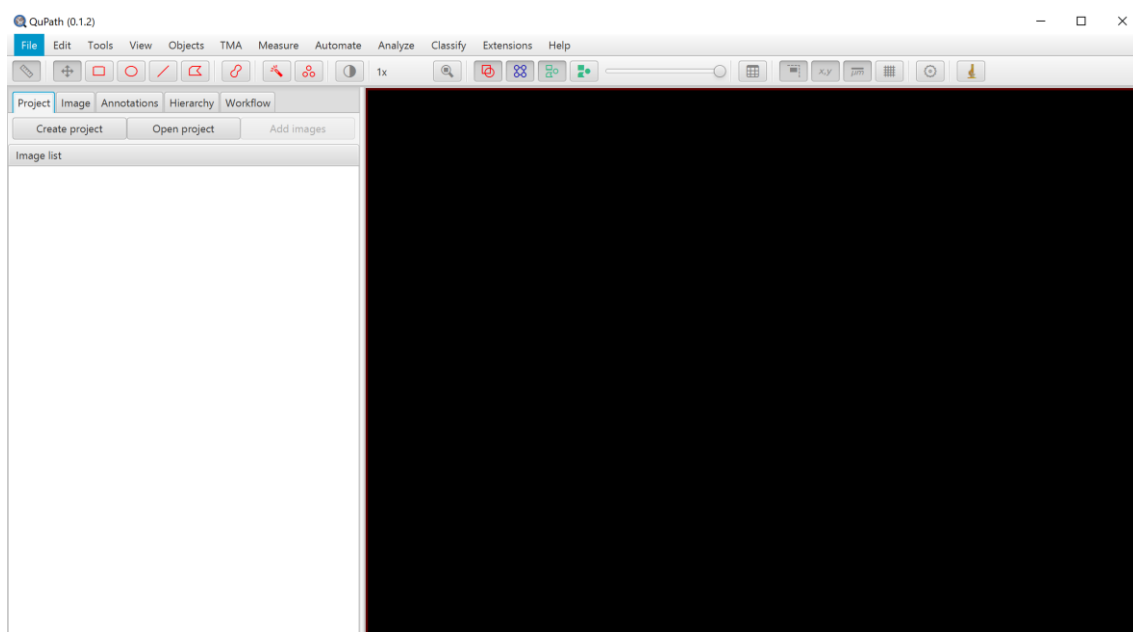
1. Setup a project for samples to be analyzed and load needed scripts
2. Run “set staining vectors” script
3. Mark representative areas from each of the samples and find proper parameters
4. Run analysis script and fine-tune parameters if needed
5. Run “results to files” and “summary table” scripts to export to Excel
6. Enjoy the results

Each of these steps are gone through in detail below.

2.2 Steps in detail

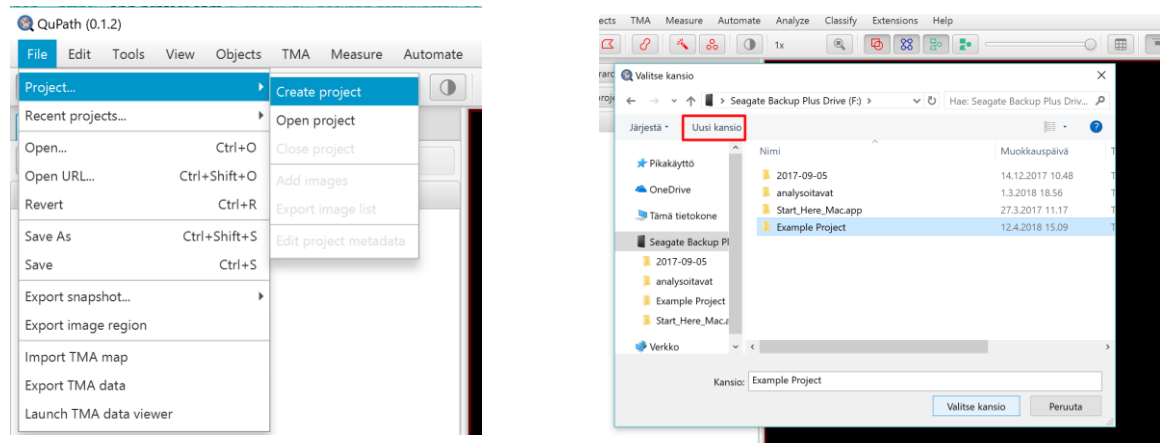
1. Setup a project for samples to be analyzed and load needed macros

Open first QuPath. The basic view should be like below:



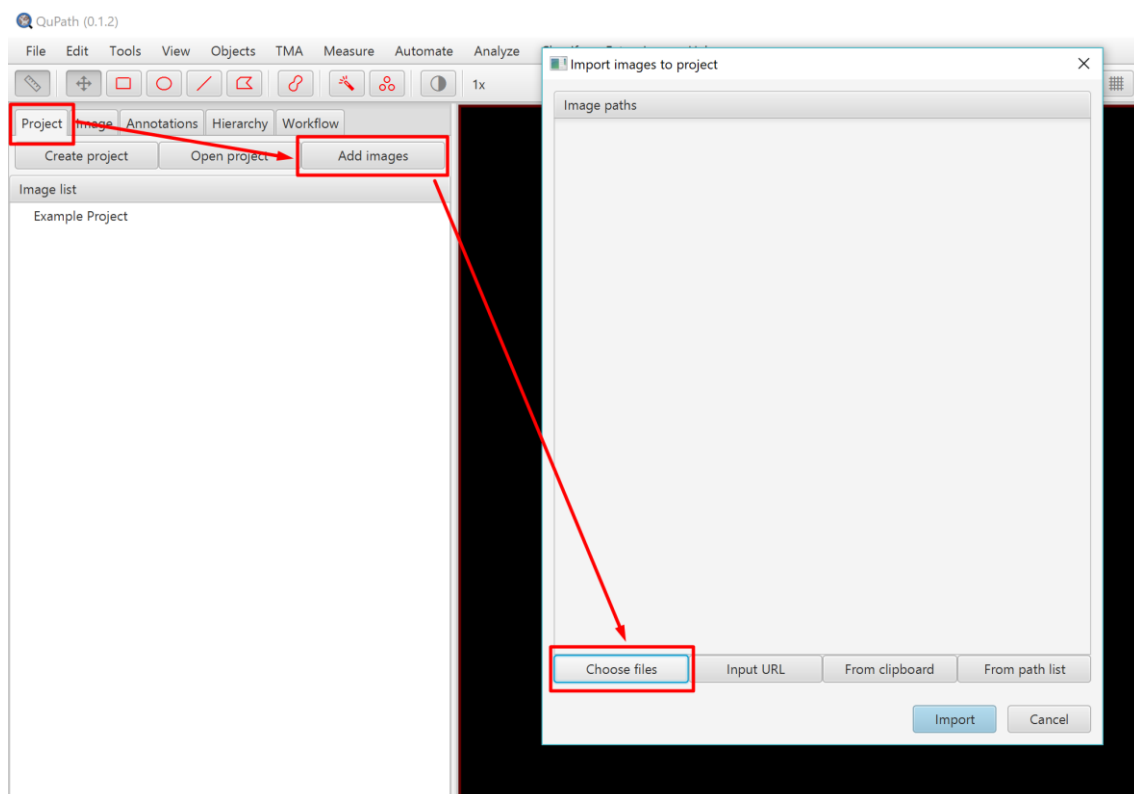
Picture 1. Basic view of QuPath

Select “Create project” and create new folder for the project. The folder can be either on same device (e.g. portable hard drive) as where the images are or on the local computer’s hard drive.



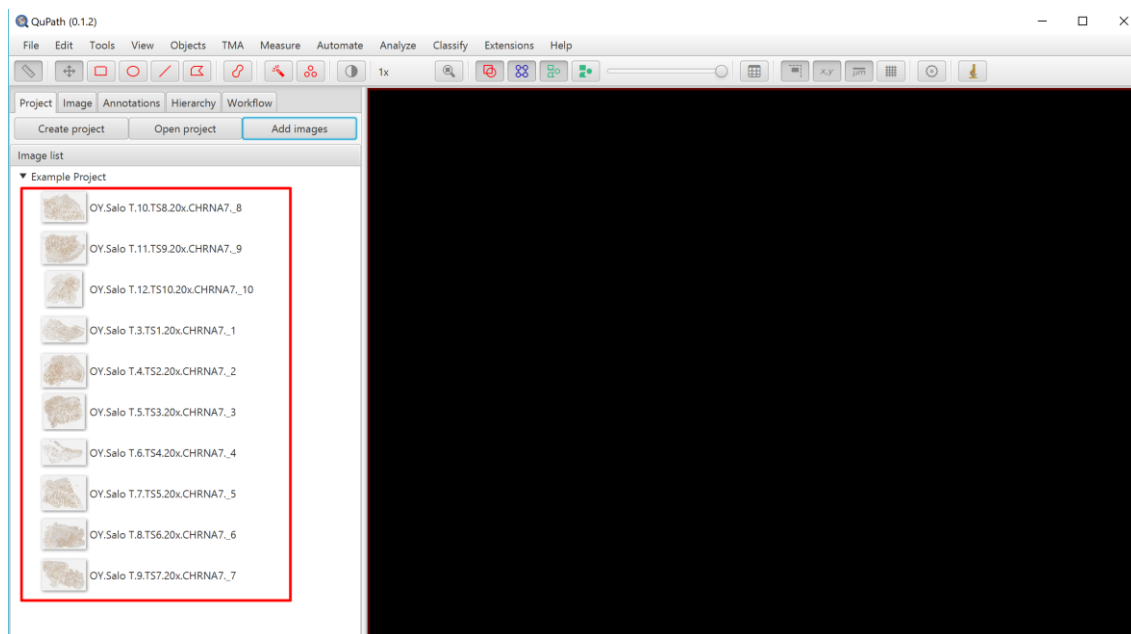
Picture 2. Creating a project and project folder

Then add wanted sample images to the project. Select “Project” tab, then “Add Images” and finally select wanted import method. For this purpose, easiest is “Choose files” after which typical file selection dialogue opens.



Picture 3. Adding images to a project

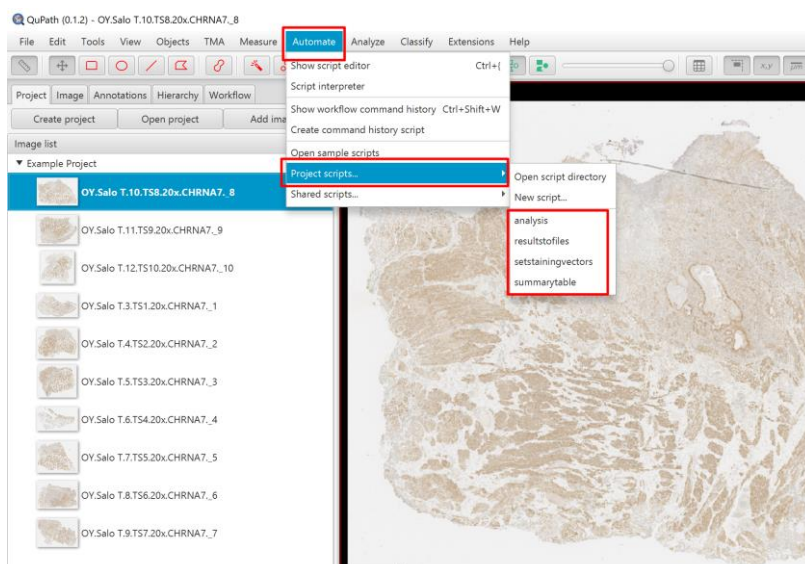
In file selection dialogue typical functions for selecting multiple files with either Ctrl- or Shift-button pressed down work and are handy. Select files and press “open”. QuPath then shows loading status bar and when completed you should see your project images as below:



Picture 4. The view of project images after adding them

Additional information: Now you can have a look at the samples by either hovering on top of the image name (small preview) or double-clicking (image will open into analysis window).

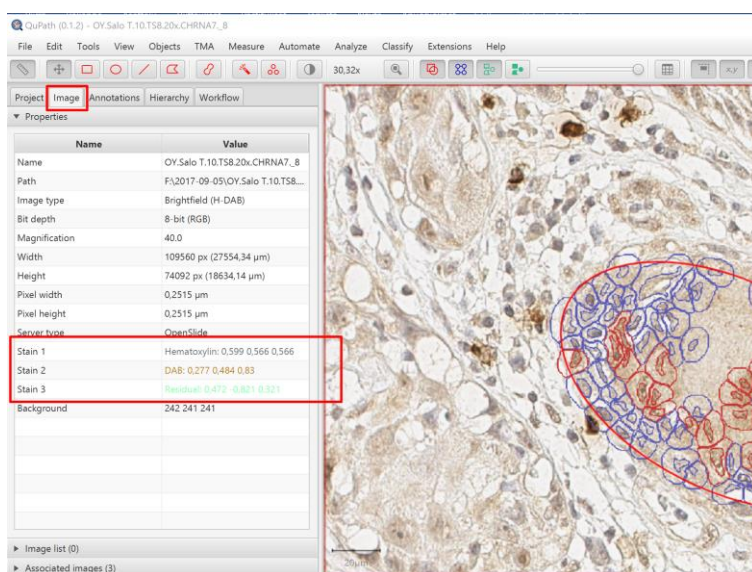
Next copy the needed scripts to the project directory. This is done outside QuPath in the normal file management of the operating system. Open the source directory for scripts to be copied (in this case: portable hard disk → “H-score scripts”). Select all the macros and copy (Ctrl-C). Go the project directory and you can see directory called “scripts” automatically created there. Go inside and paste scripts there (Ctrl-V). You can verify in QuPath that scripts are ok by hovering on “Automate” → “Project scripts” and then you should see the scripts listed in the window. No need to press anything from there at this point.



Picture 5. Script view in QuPath after copying them into a project

2. Run set staining vectors script

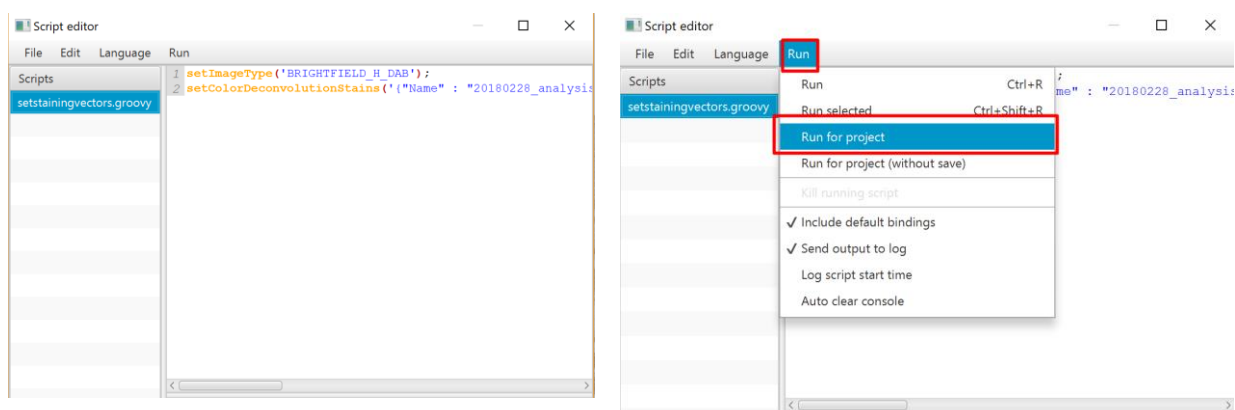
With this step the program is fine-tuned to separate stained area properly from non-stained and background areas. I.e. there are typically 3 different stain “layers” what program splits from each other. These layers’ appearance can be seen and tested simply by pressing number keys 1,2,3 and 4 respectively when one of the images is open in the analysis window. (1 = the original image, 2 = Stain 1 (in this example Hematoxylin), 3 = Stain 2 (in this example DAB), 4 = Stain 3 (in this example and typically residual, i.e. background). See picture below:



Picture 6. Staining layers’ appearance

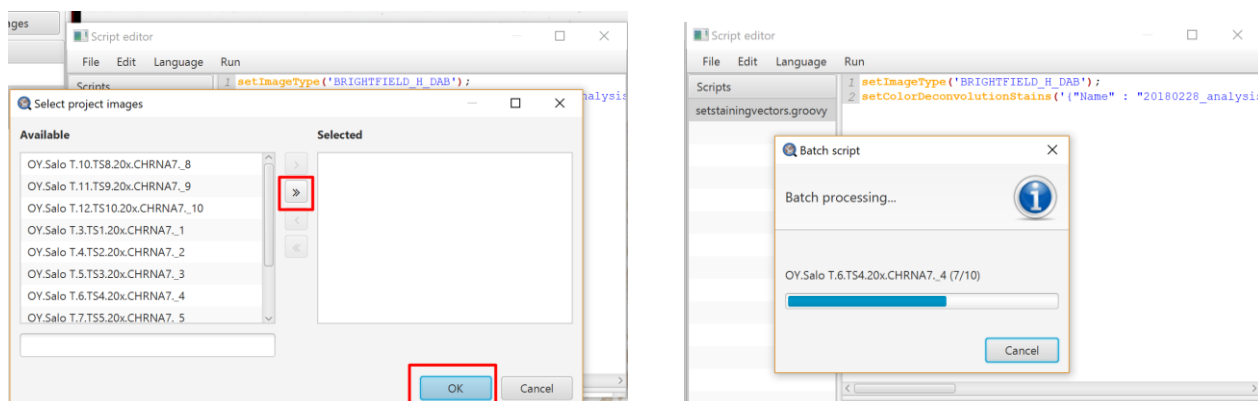
For setting the vectors simple script is ready made and needs to be run now. To be noted that the staining vectors scripts apply **only** for this image set and if another image set is used then more thorough definition of staining vectors is needed (There are instructions for this in QuPath's wiki pages: <https://github.com/qupath/qupath/wiki/Preprocessing>).

Select “Automate” → “Project scripts” → “setstainingvectors”. Then below window will open:



Picture 7. Scrip window and running a script for a project

Then “Run” → “Run for project”. This kind of dialogue box will open:

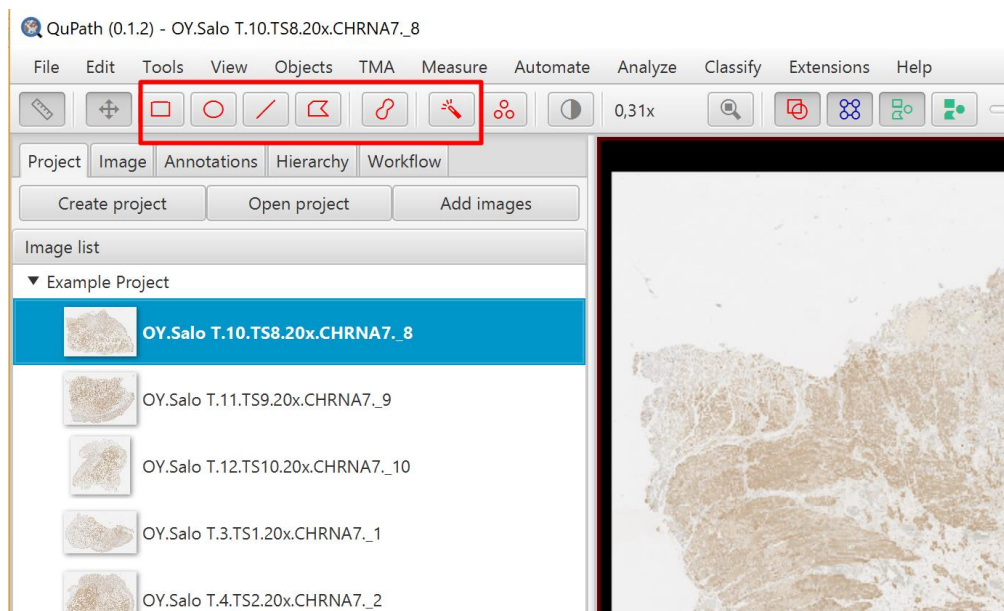


Picture 8. Dialogue for running a script for a project

This box will ask that for which images in the project current script will be driven. Select all with double arrow (then all image names will transfer to right-side box) and click “ok”. Program will show progress window like shown above. When finished Script editor –window can be closed.

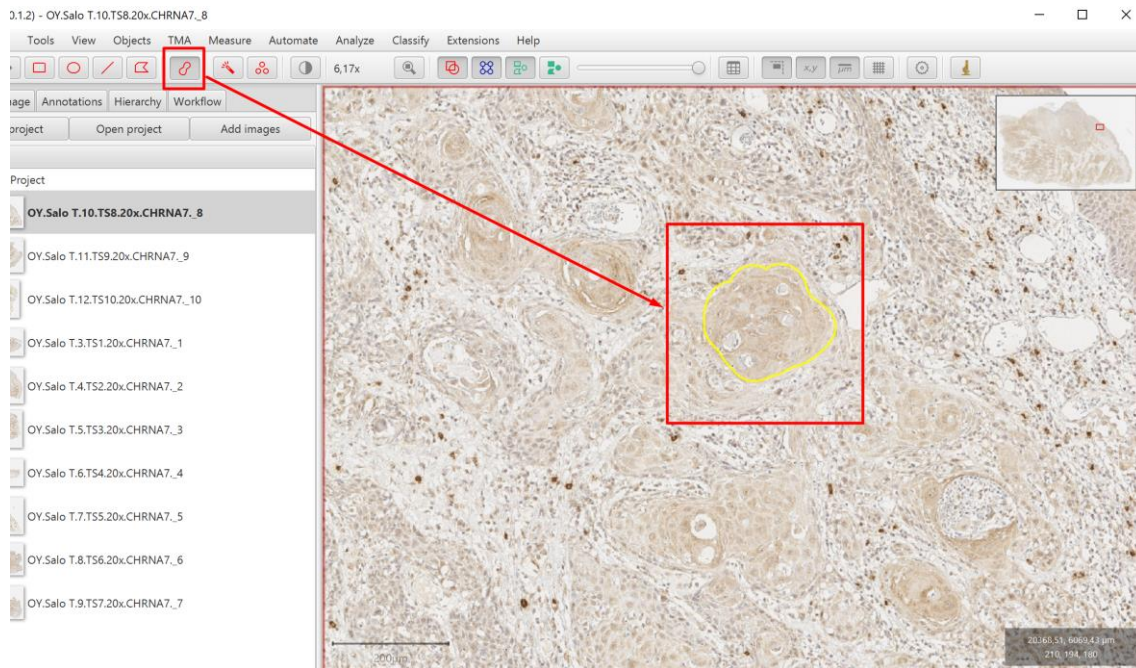
3. Mark representative areas from each of the samples and find proper parameters

Next the areas to be analyzed from each of the project images need to be selected. These areas are called *annotations*. The annotations can be made with different tools shown in picture below:



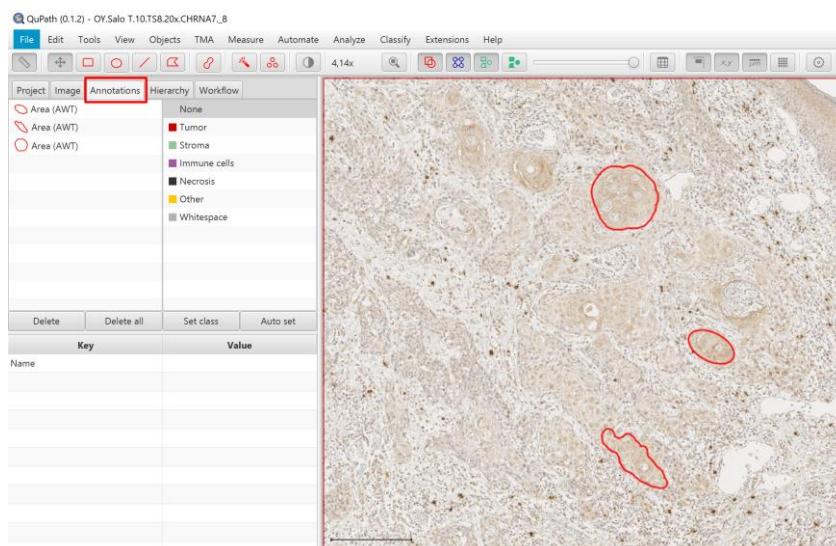
Picture 9. Tools for making annotations

Double-click the first image and it will open into analysis window. Then by zooming appropriately find the proper area of interest and mark it with one of the tools. (It is advisable to test the tools before so that you get idea that how selection works with different tools.) Example shown below:



Picture 10. Example of annotation creation

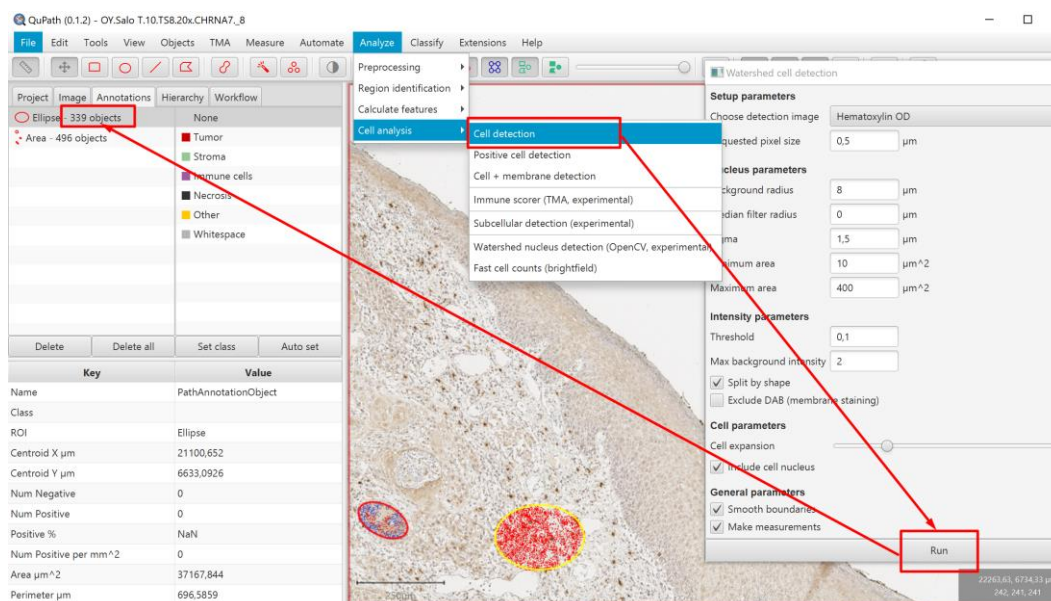
Mark in total 3 sample areas (number of areas can vary depending on the image set and agreed analysis approach). The annotations can be seen from “Annotations” –tab and can ultimately look like this for example:



Picture 11. The view of Annotations-tab and analysis window

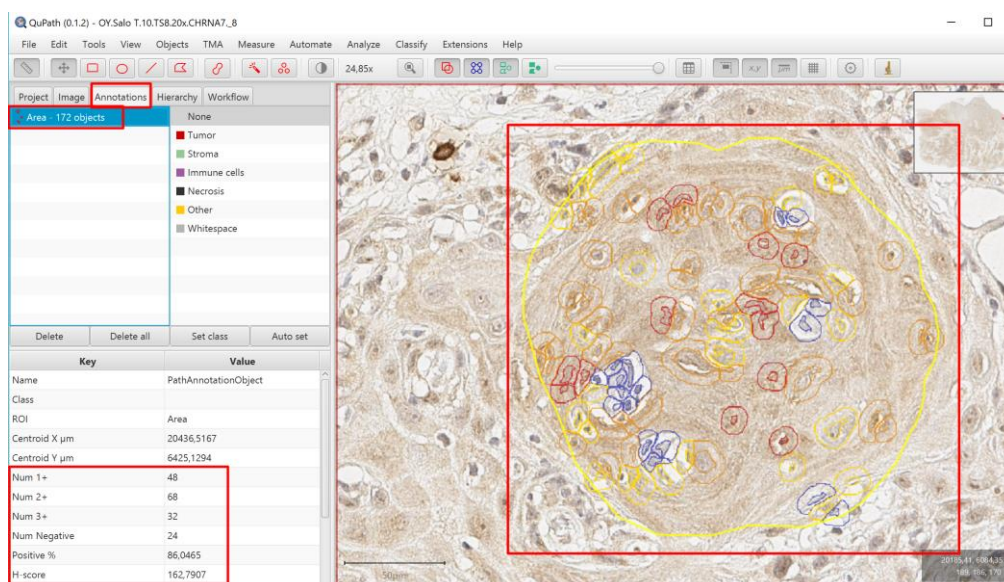
Depending on the analysis approach you may want to select sample areas which contain roughly certain number of cells. To see how many cells are in the selected annotation there is a quick method: Make sure wanted annotation is active, i.e. yellow. Then select cell analysis → cell detection → “Run”. The program will

identify and count cells and show the results next to selected annotation title (left-hand side in below screenshot). If the result is too little or too much, delete or modify the original annotation and make a new one.



Picture 12. The cell detection and cell number in the selected annotation

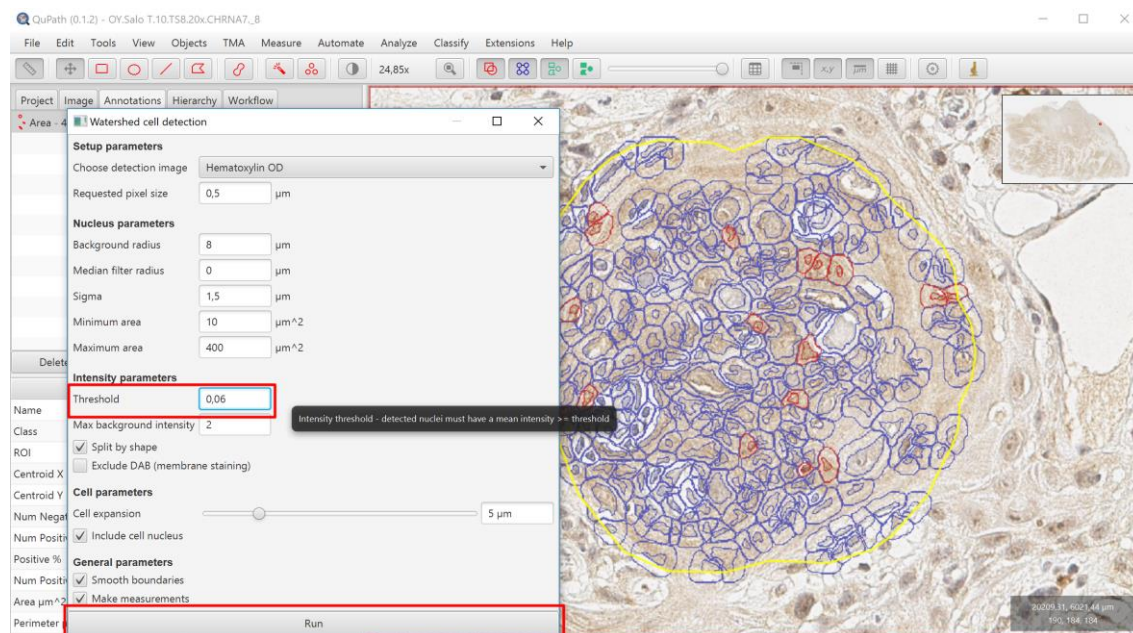
It may happen that program doesn't detect cells properly as can be seen in below example screenshot:



Picture 13. Insufficient detection of the cells

In this case it seems that detection has not worked well enough and it has missed quite a few cells. The detection can be fine-tuned by manually running simple detection which was introduced above and investigating that how detection reacts

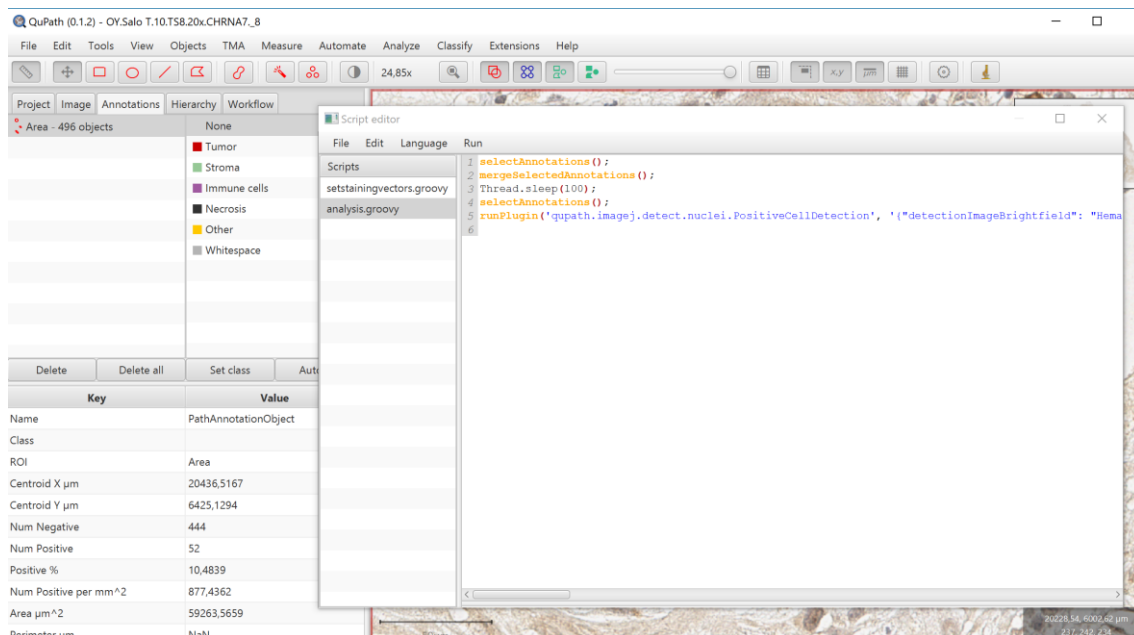
to changed parameters. Go again “Analyze” → “Cell analysis” → “Cell detection” and following dialogue will open.



Picture 14. The fine-tuning of cell detection parameters

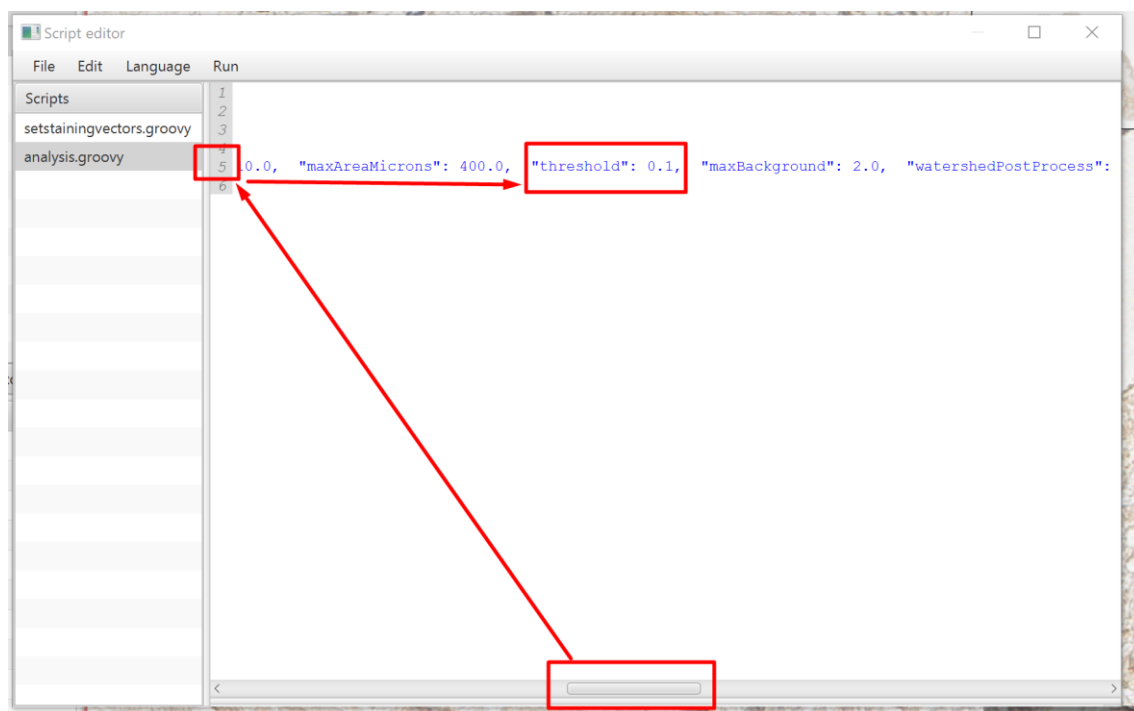
Multiple parameters can be adjusted, but for this purpose **Intensity threshold** fine-tuning can do the trick. This is as default set as 0,1. Adjust it to e.g. 0,06 and press “Run”. You can see how detection changes in the image (results shown in above screenshot’s right-hand window). After you are happy with intensity threshold value, this needs to be updated to project script so that same value is used throughout project images. Go to project scripts and “analysis” –script. Following editor view will open:

näkymä aukeaa:



Picture 15. Script editor view

From there scroll the view to right and find from row 5 the threshold parameter. Update there the value you found working with keyboard.

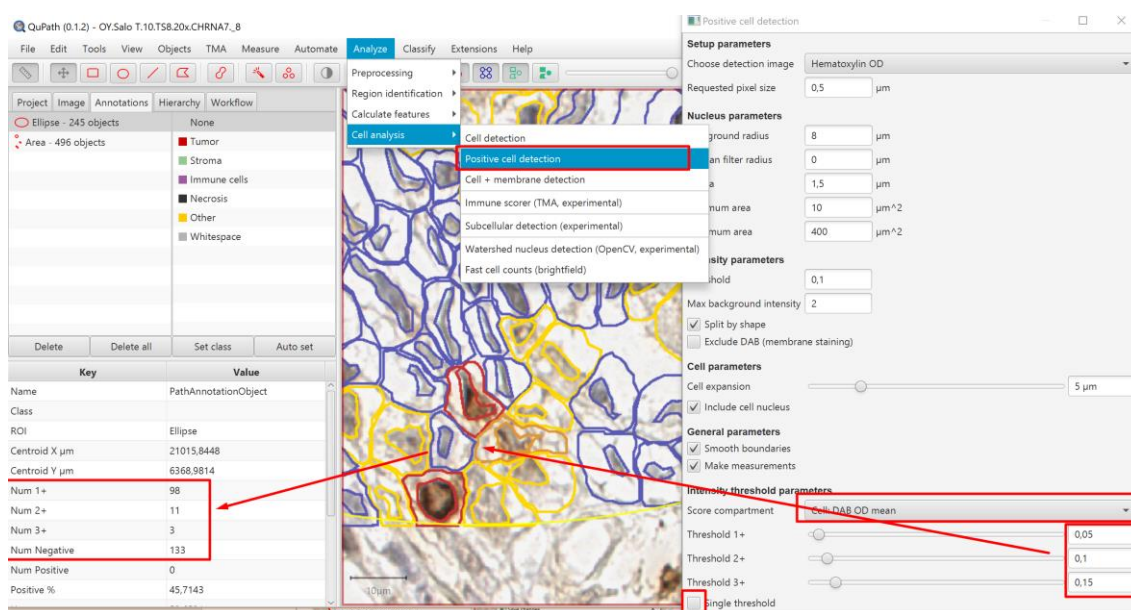


Picture 16. Finding and updating the threshold value to the script

Then in script editor select “File” → “Save”. This will save new value to script.

Other things which typically may need fine-tuning are **intensity threshold parameters**. These mean the intensity values which split the cells into intensity classes negative, class 1 (weak), class 2 (mediocre) and class 3 (strong). Select

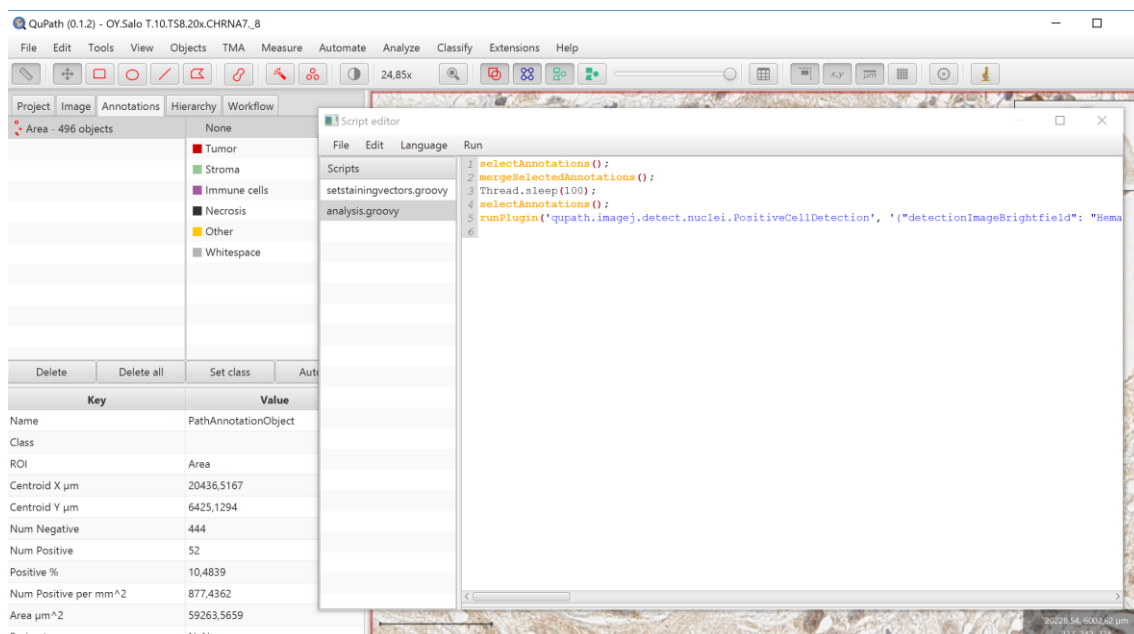
again wanted annotation and go “cell analysis” → “**positive** cell detection”. From the dialogue box certain things are now needed to change: Score compartment should be selected to match study approach and in this example it was set to include whole cell and use mean value. Then you should untick “Single threshold” -box and finally alter the threshold levels. Click “Run” and the cells get colored according to their intensity class: Blue = negative, Yellow = class 1 (weak), Orange = class 2 (mediocre), Red = class 3 (strong). The numbers of cells in each class can be also seen from the left-hand panel.



Picture 17. Altering the parameters for intensity classes

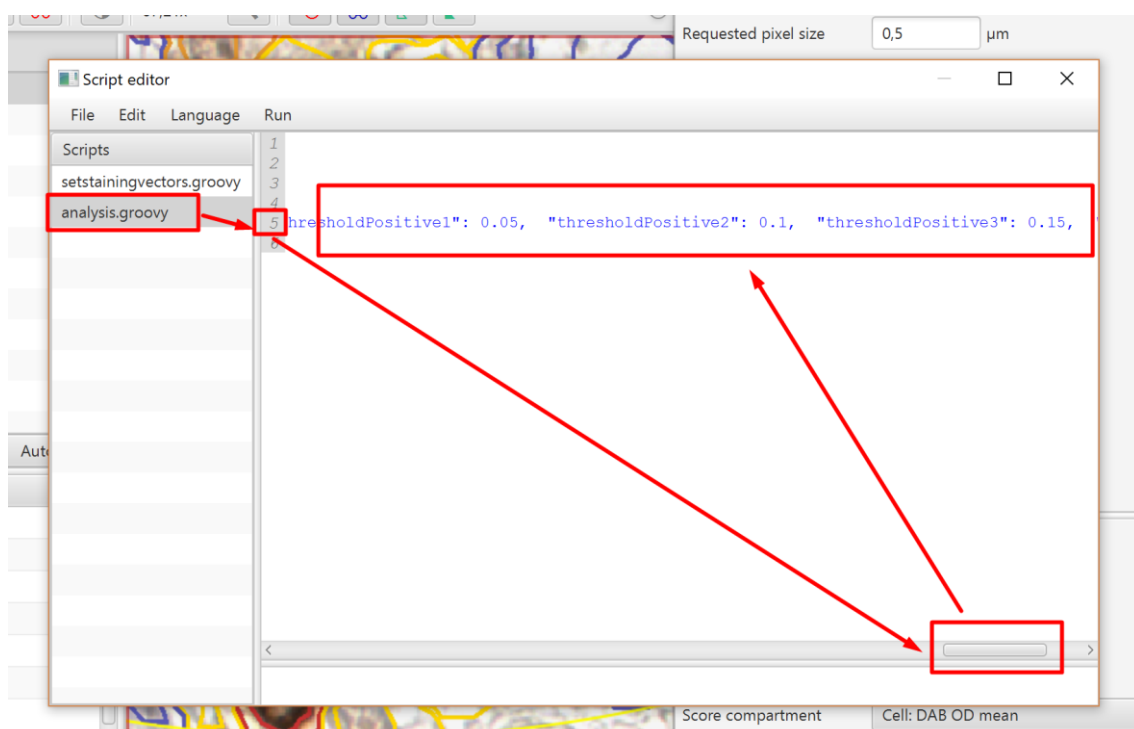
Why this is done? You may notice by eye inspection and based on your experience that thresholds need altering and again automated script needs to be updated so that new threshold values take effect to all project images. You should test with one annotation and find proper parameters and take them down.

After you are happy with intensity threshold values, these also need to be updated to project script. Go again to project scripts and “analysis” –script. Following editor view will open:



Picture 18. Script editor view for updating threshold values

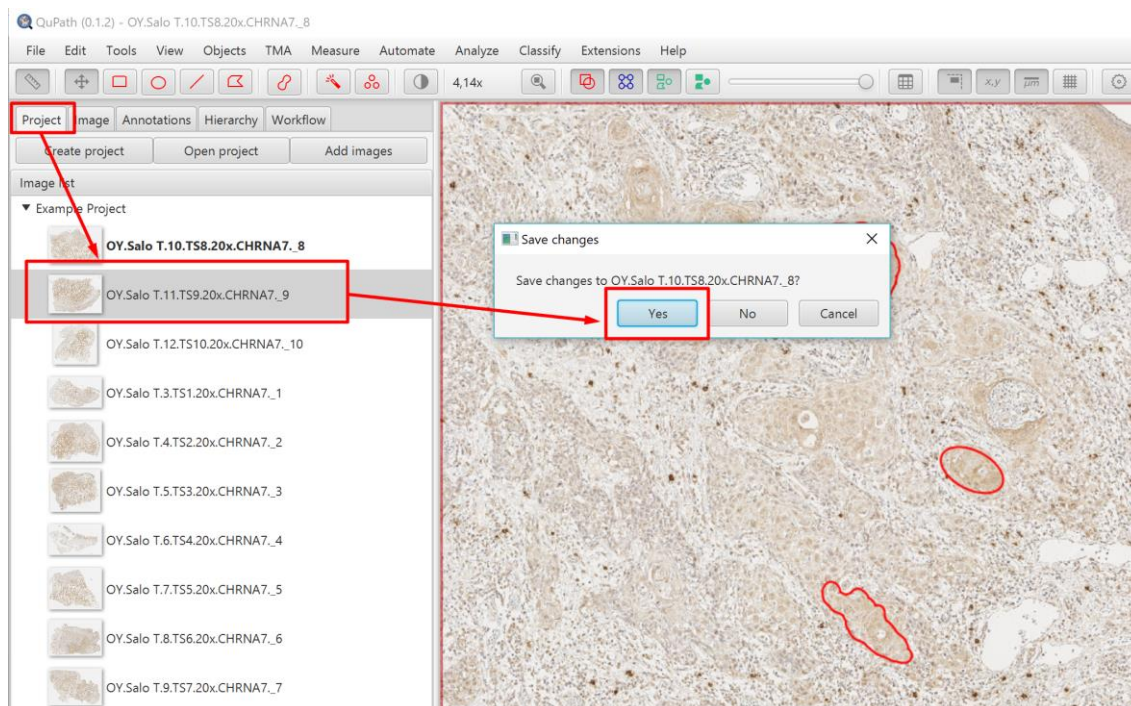
From there scroll the view to right and find from row 5 the threshold parameters. Update there the values you found working with keyboard.



Picture 19. Updating the class thresholds into the analysis script

Then in script editor select “File” → “Save”. This will save new values to script. Script editor can be closed for the timing being and you can continue making annotations.

Once you are happy with the annotations for the first image go back to “Project” –tab and select next image. The program will ask about saving and select “Yes”.



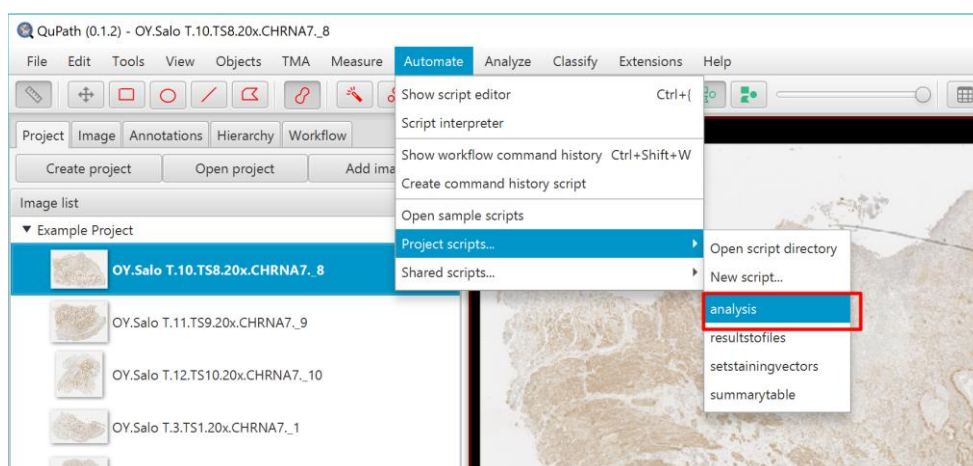
Picture 20. Proceeding in the project images and saving the changes

Repeat this area selection procedure for all the images in the project. To be noted that this is the most time-consuming step in the whole process.

Additional information: If you make an error in creating an annotation, just select the wrong annotation (it will turn yellow) and go “Objects” → “Delete” → “Delete selected objects”. Unfortunately, there is no undo-functionality in QuPath.

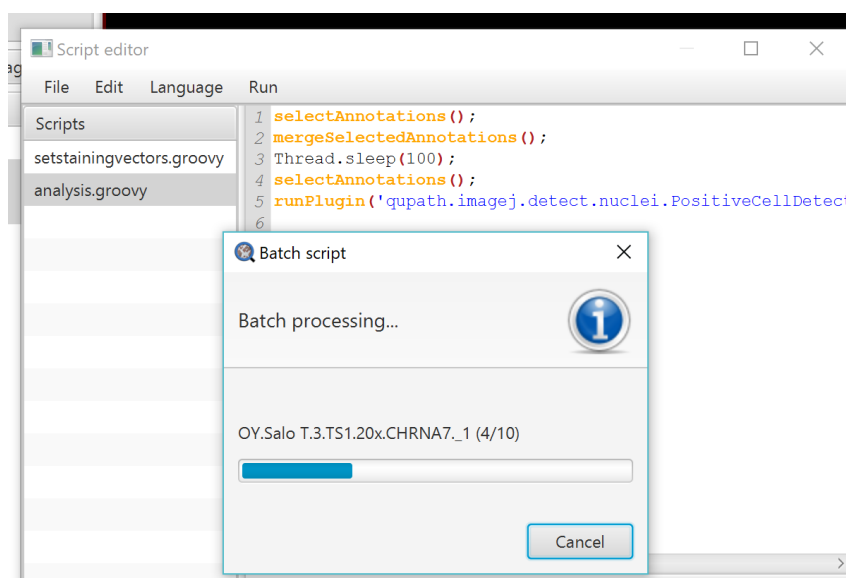
4. Run analysis script and fine-tune parameters if needed

Now all the images are saved with three annotations. Go back to first image and go “Automate” → “Project scripts” → “analysis”. This script will first combine the three annotations as one in each image and then detect and classify the cells. Select again “Run” → “Run for project”.



Picture 21. Selecting the analysis-script

Again, similar progress window will be shown:



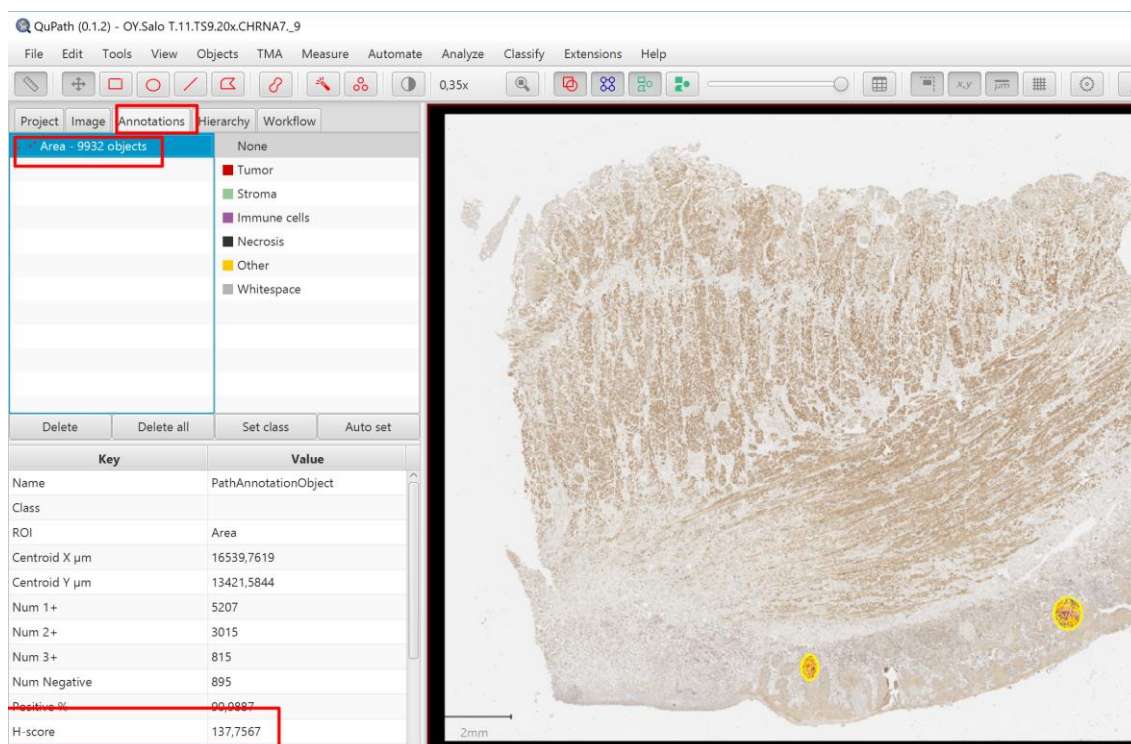
Picture 22. Progress window when script is run for the project

After finishing the script editor can be closed. The results can now be explored by double-clicking one of the images and going to “Annotations” –tab and also by zooming to annotation areas in analysis window.

If some parameter still needs changing? First it is advisable to test manually with one annotation as described below in section 3. Once new value(s) found, edit the parameter value(s) in the analysis script. Then go again “Run” and “Run for project”. This will run the analysis again to all the images and now with updated value. You can examine the results again from different images.

5. Run results to files and summary table scripts to export to Excel

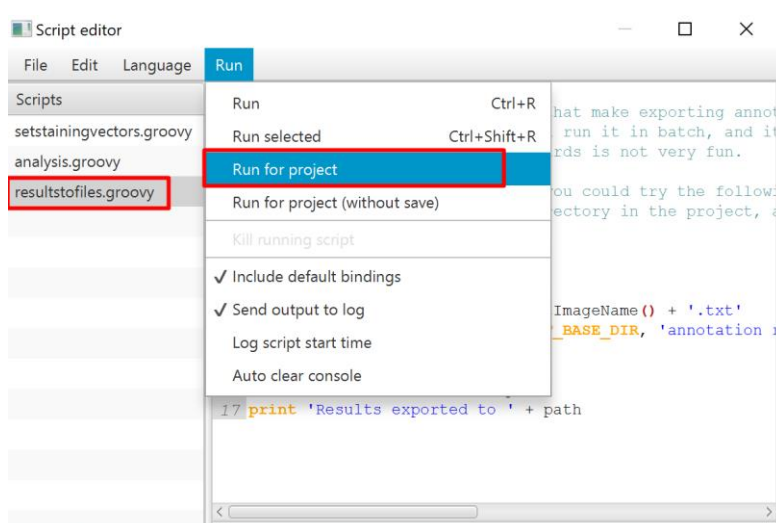
As after previous step can be seen, the H-scores are there already calculated and shown for each of the images when you select “Annotations” and select the only annotation left:



Picture 23. Viewing of H-score for selected annotation

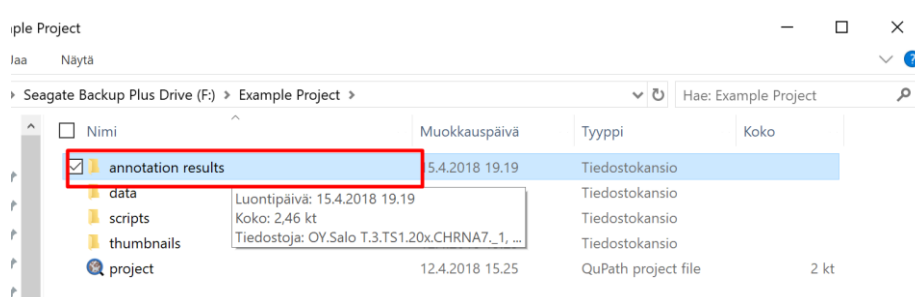
This step is meant to combine these results to one summarizing txt-file. QuPath doesn't have currently full-blown export-functionalities so therefore 2 different scripts are needed.

First run “resultstofile” script for the project:



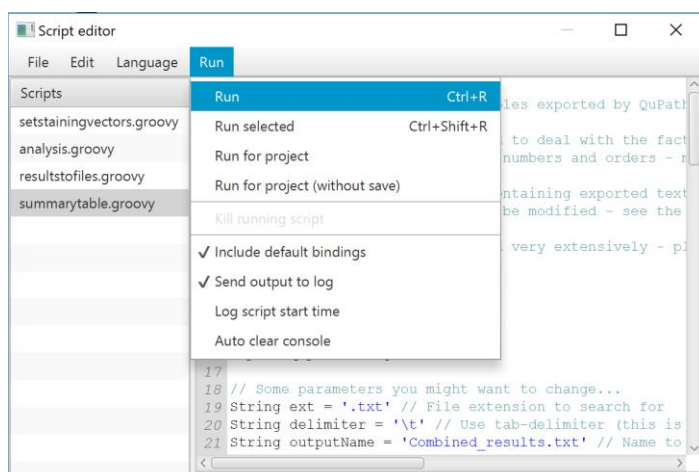
Picture 24. Running “resultstofile” script for the project

It will create “annotation results” –subfolder under your project folder.



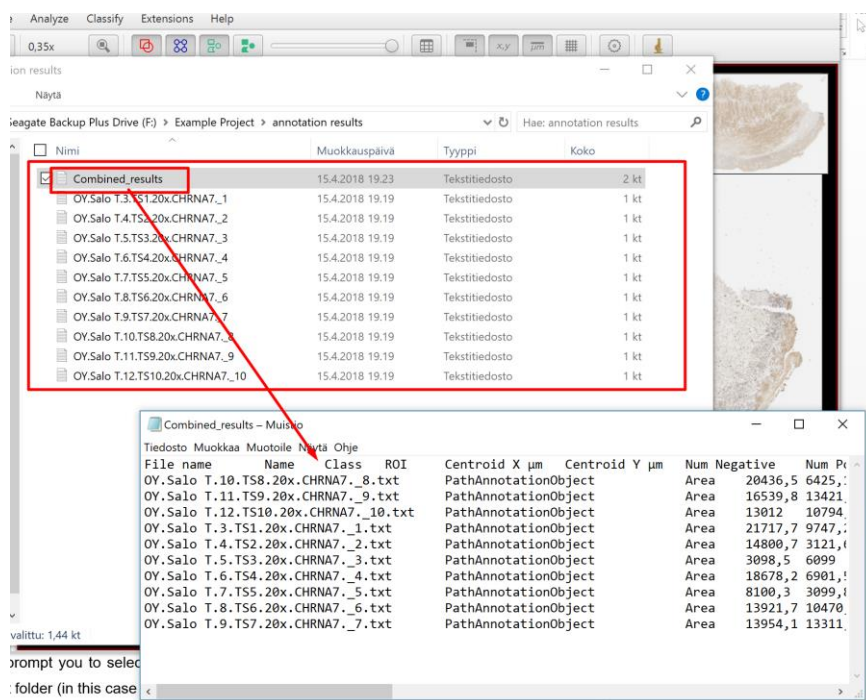
Picture 25. New subfolder under the project folder

There are now separate txt-files from each of images. Next, run “summarytable” script. NOTE that this is needed to run only as individual “Run” and no matter what image is open at background. This script will combine the earlier files to one.



Picture 26. “Run”-command for “summarytable” -script

It will prompt you to select the folder where the individual files are. Select the correct folder (in this case “annotation results”). Then you can go again in OS file management to see how folder looks. There should be now combined results – txt file and when you click it, the different image values should be there in a table format.



Picture 27. Viewing of results from the summary file

This file is better read and further analyzed by opening it in MS Excel and from there e.g. H-scores can be neatly explored:

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |
|----|-------------|----------------------|-------|-----|------------|------------|-----------|------------|------------|------------|----------|-----------|--------|--------|--------|---------|-------------|--------------|--------------|
| 1 | File name | Name | Class | ROI | Centroid X | Centroid Y | Num Negat | Num Positi | Positive % | Num Positi | Area µm² | Perimeter | Num 1+ | Num 2+ | Num 3+ | H-score | Allred prop | Allred inter | Allred score |
| 2 | OY.Salo T.1 | PathAnnotationObject | Area | | 20436,5 | 6425,1 | 444 | 52 | 10,48 | 877,44 | 59263,6 | NaN | | | | | | | |
| 3 | OY.Salo T.1 | PathAnnotationObject | Area | | 16539,8 | 13421,6 | 895 | | 90,99 | | 1267382 | NaN | 5207 | 3015 | 81 | 137,76 | 5 | 1 | 6 |
| 4 | OY.Salo T.1 | PathAnnotationObject | Area | | 13012 | 10794,2 | 1791 | | 60,38 | | 825225,9 | NaN | 1353 | 752 | 62 | 104,67 | 4 | 2 | 6 |
| 5 | OY.Salo T.3 | PathAnnotationObject | Area | | 21717,7 | 9747,2 | 188 | | 96,69 | | 535386 | NaN | 1722 | 2134 | 164 | 191,96 | 5 | 2 | 7 |
| 6 | OY.Salo T.4 | PathAnnotationObject | Area | | 14800,7 | 3121,6 | 556 | | 82,08 | | 647320 | NaN | 933 | 814 | 79 | 159,83 | 5 | 2 | 7 |
| 7 | OY.Salo T.5 | PathAnnotationObject | Area | | 3098,5 | 6099,9 | 3 | | 99,86 | | 195315,1 | NaN | 66 | 204 | 181 | 283,46 | 5 | 3 | 8 |
| 8 | OY.Salo T.6 | PathAnnotationObject | Area | | 18678,2 | 6901,5 | 1 | | 99,95 | | 221356,1 | NaN | 140 | 553 | 121 | 256,23 | 5 | 3 | 8 |
| 9 | OY.Salo T.7 | PathAnnotationObject | Area | | 8100,3 | 3099,8 | 233 | | 78,89 | | 149145,3 | NaN | 367 | 380 | 12 | 135,78 | 5 | 2 | 7 |
| 10 | OY.Salo T.8 | PathAnnotationObject | Area | | 13921,7 | 10470,8 | 18 | | 99,7 | | 529944,9 | NaN | 257 | 1704 | 396 | 261,76 | 5 | 3 | 8 |
| 11 | OY.Salo T.9 | PathAnnotationObject | Area | | 13954,1 | 13311,2 | 12 | | 97,55 | | 85112 | NaN | 51 | 86 | 34 | 254,29 | 5 | 3 | 8 |
| 12 | | | | | | | | | | | | | | | | | | | |

Picture 28. Exploring H-scores in MS Excel

6. Enjoy the results

That's it. If there are some changes needed you can go back to wanted step and do changes e.g. to cell detection parameters or areas that were manually selected.

Final note: These instructions are “just follow and trust” –type, but of course better approach would be to understand what the program and scripts do in each of the step. The program is very easy to use so learning the basics and bit more is not too time consuming. To familiarize oneself with QuPath it is strongly recommended to read wiki: <https://github.com/qupath/qupath/wiki>